DOCUMENT RESUME

ED 459 192 TM 033 478

TITLE	Measuring Up: A Standards and Assessment Benchmarking Report for Massachusetts.
INSTITUTION	Achieve, Inc., Washington, DC.
SPONS AGENCY	Massachusetts State Dept. of Education, Boston.
PUB DATE	2001-10-00
NOTE	55p.; Cover page varies. Achieve's Benchmarking Initiative.
PUB TYPE	Reports - Evaluative (142)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	*High School Students; High Schools; Mathematics Achievement; *Mathematics Instruction; Program Evaluation; *State Standards; Test Construction; *Test Use
IDENTIFIERS	*Massachusetts; Massachusetts Comprehensive Assessment System

ABSTRACT

At the request of the Massachusetts Department of Education, Achieve, Inc., conducted an evaluation of the state's K-12 mathematics standards and grade 10 Massachusetts Comprehensive Assessment System (MCAS) tests in English language arts and mathematics during the spring and summer of 2001. The state's English Language Arts standards were not analyzed because Achieve believed they were already among the best in the United States. This report presents the results of Achieve's in-depth examination of the quality, rigor, and alignment of Massachusetts' expectations. Data came from reviews by experienced reviewers. The state has made substantial progress in developing and implementing strong standards and assessments that measure what the standards expect. Overall, the evaluation finds that the standards and high school tests are of high quality and are aligned, providing a solid foundation on which to build state education policy. The grade-10 tests were found to be rigorous yet reasonable, the most challenging of the exit-level tests Achieve has reviewed. Achieve also found the mathematics standards to be well organized, jargon-free, clear, and precise. Recommendations for improvement include amendments to the MCAS high school tests to correct relatively minor flaws and development of the mathematics standards to sharpen their focus at each grade level. An appendix contains report credits and biographies of contributors. (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 459 192

EA

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
R. Schwartz
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
1

Achieve, Inc.

Measuring Up

A Report on Education Standards and Assessments for

MASSACHUSETTS

ACHIEVE'S BENCHMARKING INITIATIVE

BEST COPY AVAILABLE

TM033478

About Achieve, Inc.

Achieve is an independent, bipartisan, nonprofit organization created by governors and corporate leaders to help states and the private sector raise standards and performance in America's schools. Founded at the 1996 National Education Summit, Achieve has sponsored two additional Summits in 1999 and 2001.

Achieve helps states raise academic standards, measure performance against those standards, establish clear accountability for results and strengthen public confidence in our education system. To do this, we:

- ☐ help states **benchmark** their standards, assessments and accountability systems against the best in the country and the world;
- ☐ provide sustained **public leadership** and advocacy for the movement to raise standards and improve student performance;
- ☐ build **partnerships** that allow states to work together to improve teaching and learning and raise student achievement; and
- ☐ serve as a **national clearinghouse** on education standards and school reform.

MEASURING UP:

A STANDARDS AND ASSESSMENT BENCHMARKING REPORT FOR

MASSACHUSETTS

Prepared by Achieve, Inc. for

Commissioner David P. Driscoll
Massachusetts Department of Education

October 2001

TABLE OF CONTENTS

Executive Summary 5

Introduction: Raising Standards in America's Schools 9

The Achieve Benchmarking Methodology 13

Results for Massachusetts 19

Major Findings: Alignment of the MCAS 2001 Mathematics and English Language Arts Assessments to the Curriculum Frameworks 19

Mathematics 20

English Language Arts 27

Major Findings: Massachusetts' Mathematics Curriculum Framework 35

Recommendations for Moving Forward 45

Appendix 47

EXECUTIVE SUMMARY

Achieve Inc. was created in 1996 by governors and business leaders to serve as a clearinghouse and resource center on education standards, assessments and accountability. As part of its mission, Achieve provides states with candid feedback on the quality of their academic standards, assessments, accountability systems and other policies to promote high academic achievement. To date, Achieve has analyzed the policies and expectations of 10 states and currently is working with seven others.

At the request of the Massachusetts Department of Education, Achieve conducted an evaluation of the state's K–12 mathematics standards and grade 10 Massachusetts Comprehensive Assessment System (MCAS) tests in English language arts and mathematics during the spring and summer of 2001. The state's English language arts standards were not analyzed because Achieve believes these are already among the best standards in the nation and uses them as "exemplary standards" against which other states' standards are compared.

This report presents the results of Achieve's in-depth evaluation of the quality, rigor and alignment of Massachusetts' expectations. In particular, it provides policymakers with answers to the following questions:

- How do Massachusetts' mathematics standards compare with those of high-performing states and nations? Are the expectations for schools and students rigorous yet reasonable?

- How well do the grade 10 MCAS tests in English language arts and mathematics measure the knowledge and skills laid out in the standards? How challenging are the assessments?

RESULTS FOR MASSACHUSETTS

Massachusetts has made substantial progress in developing and implementing two essential components of standards-based reform — strong standards and assessments that measure what the standards expect. The major findings:

- **Overall, Massachusetts' standards and high school tests are of high quality and are aligned, providing a solid foundation on which to build state education policy.** The grade 10 MCAS tests are rigorous and generally well aligned with the standards, ensuring that students are required to demonstrate important knowledge and skills before graduating from high school. While the mathematics standards are not without shortcomings, which are discussed in more detail below, overall these expectations represent an articulate statement of what students should know and be able to do, at least at a minimum, by the time they complete high school. (The English language arts standards, as noted above, previously have been judged among the strongest in the country.) Importantly, this sets Massachusetts apart from

the other nine state standards and assessment programs that Achieve has reviewed — it is the only state that has both strong standards and strong assessments.

- **The grade 10 tests are rigorous yet reasonable — and are, in fact, the most challenging of the exit-level tests Achieve has reviewed.** The MCAS tests measure the important knowledge and skills demanded by the standards, are technically sound, encourage high levels of performance, and provide a template for effective classroom instruction. And releasing all common items each year is an exemplary strategy that enhances not only educational practice, but also the credibility of the state's educational improvement efforts. Many students should be able to pass these tests by the end of the 10th grade, and it is realistic to expect that the other students, given sufficient curriculum and teaching support, should be able to meet the standards by the end of high school. Students who perform below the "needs improvement" level likely have a minimal level of skill and will need intensive instructional support to achieve the minimum standards.

- **The mathematics standards generally are well organized, jargon-free, clear and precise.** The standards embody reasonable minimum-level criteria for student competency and are generally comprehensive. However, the standards should be strengthened by placing more emphasis on developing students' conceptual understanding of mathematics. Also, they are not yet as challenging as standards from Japan, Arizona or Achieve's Mathematics Achievement Partnership, of which Massachusetts is a founding partner state.

Recommendations for Moving Forward

As Massachusetts moves forward in implementing its system for standards-based educational improvement, Achieve recommends that the state consider two improvements in particular:

- ✓ **The MCAS high school tests include relatively minor flaws that should be fixed in subsequent rounds of testing. For example, the math test emphasizes some standards at the expense of others, and the English language arts test needs more treatment of nonfiction.** The MCAS mathematics test measures important content that all high school students should be responsible for knowing, yet it can be improved further. Achieve found a number of mathematics test items ostensibly designed to assess number concepts that instead more directly measure algebra standards; this has the effect of weighting the test more heavily toward algebra and omitting some advanced number concepts. Also, while the test is generally well constructed, the balance of items does not provide adequate coverage for all of the important knowledge and skills detailed in the standards. The state should ensure that items assessing numbers and data analysis are appropriately challenging on the next edition of the MCAS mathematics test.

  The language arts test is rich and rigorous, but it focuses too much on literature. All students should be exposed to a deep and engaging literary curriculum, but they also should learn how to read and interpret informational texts, such as historical documents, scientific journal

articles and technical manuals. And, the state should consider varying the 10th-grade writing prompt from year to year to assess students' skills in producing different kinds of compositions, rather than just literary analyses. By giving short shrift to students' abilities to work with informational texts, the test does not evaluate the full range of skills students need to participate meaningfully in the emerging "knowledge economy."

- ✓ **The mathematics standards should require more rigor and depth, attention to and emphasis on mathematical reasoning, and a sharper focus on essential content at each grade level.** At the middle and high school levels, students will need to be held to higher expectations to be fully prepared for success in college and high-performance workplaces. While Achieve acknowledges that many Massachusetts schools and students are struggling to meet the state's current standards — and these students will need intensive academic support in the short run — over time, as long-term changes in teacher recruitment, preparation and professional development begin to take effect, the state should raise the rigor of the mathematics standards. In the immediate future, the state may wish to publish companion materials to the mathematics standards that include numerous sample problems and activities and descriptions of how educators can build students' conceptual understanding of mathematics and reasoning ability.

* * *

Achieve also recently conducted a policy review for Massachusetts, the results of which are planned to be released in late fall of 2001. In the policy review, Achieve will address the findings of this report in addition to other core elements of a comprehensive system of standards-based education, such as capacity building, accountability and public engagement. By taking a hard look at the progress that has been made in the more than eight years Massachusetts has invested in implementing standards-based systems — and by identifying important work still to be done — Achieve hopes to help the state meet its goal of raising the achievement of all its students.

INTRODUCTION: RAISING STANDARDS IN AMERICA'S SCHOOLS

Since the release of *A Nation at Risk* in 1983, schools, states and national policymakers have been concerned with improving the level of academic achievement of all students. At that time, most school systems awarded diplomas based on Carnegie units, which generally represented "seat time," as opposed to a demonstration of knowledge and skills. Under the Carnegie or credit-based system, postsecondary institutions and employers had difficulty determining what students had learned, since course content varied from school to school. Additionally, it was increasingly clear that a large percentage of our nation's students, particularly minorities and the poor, were being dramatically underserved by their education system because they were not provided with the same rich curriculum and learning opportunities as their wealthier counterparts in the suburbs.

In an attempt to raise the level of student learning across the board and create a more publicly accountable education system, states, districts and national organizations began the process of drafting content standards to define what students should know and be able to do. In Massachusetts and across the country, standards are now the driving force in efforts to improve equity and excellence in education by holding *all* students to common, high expectations.

Since the early 1990s, 49 states have developed academic standards for their students, and 48 are putting in place assessments to measure those standards. By stating clearly the knowledge and skills students are expected to gain as a result of their schooling, reformers hope that students will better understand what is expected of them, schools will improve their programs to help students achieve those expectations, and low-performing districts and schools in particular will be challenged to raise the level of teaching and learning. Because states have made substantial investments in the new standards and tests — and many states are beginning to hold students and schools accountable for performance — policymakers and the public want to know how their standards compare to what other states and countries expect. Achieve was created precisely to address this issue.

Born out of the 1996 National Education Summit, Achieve helps states ensure that they have in place standards that compare favorably with the academic expectations of other states and high-performing nations and assessments that accurately measure student achievement against those standards. An independent, bipartisan, nonprofit organization overseen by a board of directors composed of governors and corporate CEOs, Achieve serves as a clearinghouse and resource center on education standards, testing and accountability, working primarily with states to support their work in these areas.

We have come a long way in the effort to improve schools for all students through the development and implementation of standards-based reform. In the initial stage of the standards movement, states found it challenging to develop quality standards. We now have a better picture of what strong academic standards look like. They are clear and specific enough to guide curriculum planning and test development without infringing on local control. They set rigorous yet reasonable expectations for all students and raise the bar higher than it is currently set for

many students. They integrate content knowledge with important thinking skills and learning processes. And they are widely read and understood by parents, educators, business people and policymakers. State assessments are evolving as well. More attention is being given to alignment, to incorporating a mix of formats (multiple choice, short answer and open-ended items) and to sharing concrete information with district educators and the general public.

States are also revisiting the issue of local control, with each state striking a different balance. States are realizing that respecting local control while ensuring high standards for all students means being explicit about the knowledge and skills students are expected to learn and assisting educators by providing tools and training. This new understanding is at the heart of standards-based reform: The previous state role of monitoring compliance to rules and regulations has shifted to one of setting expectations for results and supporting schools and districts in meeting those expectations.

States and schools are currently grappling with a host of thorny issues that accompany full implementation of standards-based reform, including:

- strengthening teacher certification and professional development;
- developing or identifying materials to support state standards and assessments;
- setting fair and defensible promotion and graduation requirements;
- providing additional funding and programs for struggling students;
- ensuring equitable treatment of special populations; and
- establishing genuine accountability for all education stakeholders and policymakers.

In a separate, yet complementary, report, "Taking Stock," Achieve will examine the progress Massachusetts has made to date in moving toward full implementation of a standards-based system. Achieve's report, slated for completion in late fall 2001, will focus on issues of equity and excellence and make recommendations to the state for the next decade of reform.

Benchmarking to the Best

To help states like Massachusetts in their efforts to continuously improve expectations for all students, Achieve provides *standards and assessment benchmarking*. Through benchmarking, Achieve compares a state's academic expectations against the best available models from the United States and the world. States like Massachusetts that have sought benchmarking services from Achieve are committed to raising standards for student performance and holding schools accountable. These states want their citizens to know that the standards they have set compare favorably with the expectations other states and nations have for their students. They also want to understand whether the tests states use to assess student progress against the standards truly measure what they expect all students to know and be able to do. Lastly, they want objective, credible, concrete recommendations for ways to improve their standards and assessments.

Benchmarking is a highly respected practice in the business world. It is an activity that looks outward to find best practices and high performance and then measures actual business operations against those goals. Benchmarking in education follows the same principle. It is appropriate at a time when state education reforms are focused on raising student and school performance, as states want and need an external yardstick to gauge their efforts.

By benchmarking academic standards and assessments, Achieve hopes to help states answer the following questions:

- How do our education standards compare with those of other high-performing states and nations? Are the expectations for our students and schools high enough?

- How well do our assessments measure the knowledge and skills laid out in the standards?

Achieve is involved in benchmarking for another important reason: States have traditionally had limited access to high-quality, trustworthy information about education standards. This is due partly to the fact that the standards movement in education is relatively young. But it is also a result of the disparate nature of much of the work that has been done to date. While the standards reviews and "report cards" issued by other organizations have helped to focus national attention on the quality of standards, their judgments have often been in conflict, and their tone has not always been constructive. States increasingly are looking for independent, credible advice on these issues.

Achieve's benchmarking efforts are not designed to grade or rank states. Instead, we have created a service that is diagnostic in nature — one that yields detailed, reliable information that we hope states will find useful. In addition, our focus on assessments as well as standards and the alignment between the two allows us to truly determine what the state expects all its students to know and be able to do — and whether the standards and assessments are a strong enough foundation for the state's efforts to improve education performance.

THE ACHIEVE BENCHMARKING METHODOLOGY

Achieve's benchmarking methodology has been developed and tested over four years. Massachusetts is the tenth state to put itself through this intensive process. Achieve staff and consultants have reviewed dozens of standards and tests and bring their experience to bear in applying the benchmarking tools to Massachusetts' standards and tests.

STANDARDS BENCHMARKING

Achieve compares a state's standards to state, national and international benchmark standards recognized for their quality and/or for producing high student achievement. This comparison of state standards to benchmark standards is designed to answer the following questions:

- Do the standards define a comprehensive, yet manageable academic core for all students? Are there key concepts or skills missing in the standards?

- Do the standards convey both the breadth and depth of knowledge and skills expected of students? Have choices been made about what is most important for students to learn, and when?

- Are the standards as rigorous as they should be? Do the expectations described by the state standards compare favorably to those of the benchmark standards?

- Do the standards define both what students should know (i.e., content knowledge) and what they should be able to do with that knowledge? Are reasoning and problem solving skills fully developed? Or is one overemphasized at the expense of the other?

- Are the standards clear and explicit, conveyed in a way that educators and parents can understand and use to improve student achievement?

To ensure that the benchmark standards documents used as exemplars are indeed the best for this purpose, in early 1999, Achieve commissioned expert reviews of a variety of sets of standards. Reviewers concluded that California's and Massachusetts' standards in English, those of North Carolina, Texas and New Standards in early literacy, and those of Arizona and Japan in mathematics had the most value for benchmarking.[1]

Selecting these benchmarks proved to be a difficult task because no one set of standards is perfect, and judgments about the quality of standards are in some ways subjective. Still, we are confident that the choices used in our current work reflect some of the best thinking from around the country.

[1] Achieve created "benchmark profiles" for each of these documents that provide contextual information about the standards and summarize their strengths and weaknesses. The profiles are available upon request.

and that a careful comparison of the state standards to these benchmarks will yield helpful diagnostic information and policy suggestions for states to consider.

For example, in the area of language arts, California distinguishes those standards concerned with informational text from those concerned with literature. This makes it easy for teachers to help students grasp the different strategies used in reading, understanding and writing the two types of text. Massachusetts organizes its standards in grade spans of two years, as do a number of states, but Massachusetts articulates its expectations for student learning more precisely than do most states. In mathematics, Arizona's standards provide a level of specificity and detail that is helpful for teachers — especially for elementary generalists — while Japan's standards provide an economical, yet focused and rigorous treatment of the discipline.

Assessment-to-Standards Alignment Analysis

Achieve's assessment-to-standards analysis is designed to address the alignment of tests to standards. It helps uncover answers to the following issues:

- **Fairness.** Does each assessment only measure content and skills reflected in the standards? Or put differently, can everything on the test be found in the state standards?

- **Balance.** Does each assessment measure the breadth and depth of content and skill in the standards? In other words, to what extent does each assessment measure the key content and skills for a grade level?

- **Rigor.** Overall, is each assessment sufficiently challenging for students? Do the assessments grow more sophisticated from grade to grade?

Alignment is not a "yes or no" question — nor it is a mathematical calculation. It is the extent to which standards and assessments are in agreement and serve in conjunction with one another to guide and support student learning. Consequently, responding to the above questions requires a systematic procedure to probe the different factors that contribute to alignment. Achieve's process or "protocol" for determining the alignment of assessments to standards is based upon five criteria. Application of the protocol provides rich information about alignment of tests and standards, which is typically unavailable to states. The fundamental criteria for Achieve's alignment process are:

- **Confirmation or construction of test blueprint**. Reviewers check to see that each item corresponds to at least one standard or objective. If no test blueprint is provided by the state or testing company, then Achieve's reviewers construct a new one. If the blueprint supplied by the test developer does not stand up to scrutiny (i.e., a significant number of items mapped to one standard or objective are found to be more closely related to a different one), reviewers make this known to the state.

- **Content centrality**. This criterion examines the quality of the match between the content of each test item and the content of the related standard. Reviewers determine how closely the content of the item aligns with that of the related standard, and then assign the item to one of four categories based on degree of alignment, from "not aligned" to "clearly aligned."

- **Performance centrality**. Each item places a certain type of cognitive demand on a student (e.g., the student is asked to "identify" or "analyze"). If an item simply requires a student to "identify" and the corresponding standard requires a student to "analyze," then there is a mismatch between the two performances. Reviewers assign each item to one of four categories based on the degree of alignment, from "not aligned" to "clearly aligned."

- **Challenge**. This criterion is applied to both the *individual items* and to the *set of items* that measure an entire strand, such as "Measurement." Its purpose is to determine whether doing well on these items requires students to master challenging subject matter. At the item level, reviewers consider two factors related to challenge: *source of challenge* and *level of cognitive demand*. At the set level, reviewers consider the overall *level of challenge* of the items mapped to a strand.

    - *Source of challenge*. This criterion attempts to uncover whether an individual test item is "fair." Reviewers analyze whether an item is difficult because of the knowledge and skills it targets, or for other reasons not related to the subject matter, such as relying unfairly on students' background knowledge, and rate each item as having an appropriate or inappropriate source of challenge. Any item judged to have an inappropriate source of challenge is *not* examined when item sets are evaluated for level of challenge.

    - *Level of cognitive demand*. This criterion focuses on the type and level of thinking and reasoning required by the student on a particular item.[2] A Level 1 (recall) item requires the recall of information such as a fact, definition, term or simple procedure. A Level 2 (skill/concept) item calls for the engagement of some mental processing beyond a habitual response, with students required to make some decisions as to how to approach a problem or activity. Level 3 (strategic thinking) items require students to reason, plan or use evidence. And Level 4 (extended thinking) items require complex reasoning, planning, developing and thinking, typically over an extended period of time.

    - *Level of challenge*. This term applies to the set of items that maps to a standard. Reviewers compare the overall demand encompassed by a set of items to the level of demand expressed in the standard itself. In addition to evaluating alignment, reviewers also judge whether the set of test items has a span of difficulty appropriate for students

[2] Norman L. Webb. (2001) *Levels for Determining Depth of Knowledge*. CCSSO TILSA Alignment Study, Version 2.0, May 21-24, 2001.

at a given grade level. This judgment is based on the standards, the assessment and supporting materials such as student responses. Reviewers write a succinct summary of the level of challenge of each item set.

- **Balance and range**. Balance compares the extent to which the knowledge and skills delineated in the standards receive the same emphasis on the assessment, and if that emphasis is appropriate. Range examines the degree to which the assessments sample the knowledge and skills described in the standards, since it is very difficult for one assessment to measure the full complement of knowledge and skills required by state standards. Evaluating balance and range provides both qualitative and quantitative information about the choices states or test developers have made.

Working With Massachusetts

The Achieve benchmarking process examined Massachusetts' mathematics standards against our benchmark standards and then compared Massachusetts' grade 10 MCAS 2001 test in mathematics and English language arts against the state standards.

Achieve did not benchmark Massachusetts' standards in English language arts — in effect, that judgment has already been made. As noted previously, when Achieve conducted its original search for exemplars, the 1997 Massachusetts Framework in English language arts emerged as one of the best standards documents in the country. The newly revised 2000 Framework contains significant improvements and has replaced the former edition as one of Achieve's benchmarks in English language arts.

As the first step in analyzing Massachusetts' mathematics standards, Achieve senior staff systematically compared the content and skills found in the Massachusetts standards to those of Arizona, Japan and Achieve's MAP expectations. Five national experts in standards then reviewed Massachusetts' standards and the comparisons to the benchmark standards and responded to a set of guiding questions. These experts hold diverse opinions about subject matter, curriculum and assessment issues, and each has considerable experience in writing, researching and analyzing standards.

Achieve's alignment analysis is a process of managing expert judgment. There is no mathematical formula for matching a test to standards. Rather, the process relies on experienced, knowledgeable educators who bring their experience and knowledge to bear in applying the criteria for judging alignment.

To gauge the alignment of Massachusetts' grade 10 tests to the standards, Achieve convened two teams of highly skilled educators to carefully study the tests and apply the Achieve alignment protocol. The reviewers selected by Achieve to analyze the alignment between Massachusetts' assessments and standards are a deliberate mix of classroom teachers, curriculum specialists and subject-matter experts, each with extensive expertise in content and assessment design. They often

have experience in large-scale assessments and/or standards development, represent a diversity of viewpoints on curriculum matters, and have worked in a variety of challenging school environments in rural, suburban and urban settings.

This diversity of backgrounds has proven invaluable in arriving at considered judgments (for example, determining the appropriateness of a test item for a particular grade level). Achieve's senior consultants in mathematics and English language arts led their teams through a stepwise application of the protocol, with the goal of arriving at consensus judgments for each of the alignment criteria.

In writing this report, Achieve synthesized the five reviews of the Massachusetts mathematics standards, highlighting the strengths and weaknesses of the document as identified by the subject area experts, and summarized the results of the alignment studies as reported by the review teams in mathematics and English language arts. The findings described in this report represent consensus opinions of Achieve's consultants and experts, but final judgments and conclusions rest with Achieve.

Brief biographies of Achieve's experts and consultants who participated in the standards benchmarking and assessment analysis for Massachusetts can be found in the Appendix.

RESULTS FOR MASSACHUSETTS

MAJOR FINDINGS: ALIGNMENT OF THE MCAS 2001 MATHEMATICS AND ENGLISH LANGUAGE ARTS ASSESSMENTS TO THE CURRICULUM FRAMEWORKS

Discipline-based teams of experienced reviewers examined the alignment of the grade 10 MCAS 2001 Mathematics test to the standards for grades 9–10 in the 2000 Mathematics Curriculum Framework and the alignment of the grade 10 MCAS English language arts test to the 9–12 standards in the 1997 English language arts Curriculum Framework, using Achieve's assessment-to-standards alignment protocol. Brief biographies of Achieve's expert reviewers are included in the appendix.

The Grade 10 MCAS mathematics test consists of three types of items: multiple choice, short answer and extended open response. The grade 10 English language arts test also consists of three item types: multiple choice, extended open response and a writing prompt. Short answer and open response questions make up 40 percent of the total possible score on the mathematics test (60 points maximum). The writing prompt and open response questions make up about 50 percent of the total possible score (72) on the English language arts test.

There are multiple forms for both of the mathematics and English assessments. Approximately 80 percent of the items in any form are identical. These items are referred to as "common" questions. The remaining 20 percent of the items vary from form to form and are called "matrixed-sampled" items. Student, school and district scores are based exclusively on common items. The common items are released after each administration of the test and replaced with items from the pool of matrixed-sampled items. A portion of the matrixed-sampled items are not used as replacement items, but are repeated across test administrations and used to link (equate) the tests from year to year. Because only common items are used to determine student, school and district scores, Achieve's reviewers were asked to limit their analysis to these questions.

The MCAS grade 10 mathematics test is administered in three sessions, with students being allowed to use calculators during two of the three sessions. The MCAS Grade 10 English Language Arts test contains two distinct parts: The Composition portion is based on a writing prompt, is structured to include some of the key elements of the writing process (drafting, revising and finalizing) and is administered in two sessions. The Language and Literature portion is based on grade appropriate reading passages, both fiction and nonfiction, and is administered in three test sessions. In English, students may use dictionaries for the writing prompt, and students with limited English proficiency can use bilingual word-to-word dictionaries; in mathematics, a Formula Reference Sheet is provided. MCAS tests are not strictly timed, and the state has made provisions for schools to extend the recommended testing time of 45 minutes per session.

MATHEMATICS

The 2000 Mathematics Curriculum Framework is organized into five strands (Number Sense and Operations; Patterns, Relations and Algebra; Geometry; Measurement; and Data Analysis, Statistics and Probability), whereas the 1996 framework contained four strands (Geometry and Measurement were combined). Massachusetts' decision to report 2001 results using the four 1996 strands, and to base test scores solely on learning standards common to both the 1996 and 2000 frameworks, will allow the state to ensure that scores at student, school and district levels are fair.

Strengths of the Grade 10 Assessment

Achieve's reviewers identified the following strengths in the grade 10 MCAS 2001 mathematics test:

1. **Massachusetts is committed to sharing assessment information with all stakeholders; what is tested is not a state secret.**

In its February 2001 document, *Overview of the MCAS 2001 Tests,* the state presented the distribution of item types and points for the four mathematics strands. Overall, the 2001 mathematics test adheres closely to those descriptions. Moreover, the percent of total points allocated to each item type is fair and appropriate across the test as a whole. This public presentation of the test elements is not common to all state assessment systems and is an important part of the mutual accountability effort displayed by Massachusetts.

Reviewers were impressed with the clarity and amount of information provided by the state on its assessments and commend the state for providing such easy access to information. The detailed feedback provided by MCAS reports at the student, school and district levels helps ensure that the testing is beneficial for individual diagnosis and curriculum and instructional planning. Massachusetts is also to be commended for the release of its common test items so promptly after test administration. The release of these items helps ensure public credibility and provides concrete models for teachers to use as they design and develop instructional activities. Such a high degree of public accessibility to information about a state assessment system is rare, yet essential to sustaining public support for higher standards.

2. **The test assesses only content and skills included in the standards, and test items generally align — either in their entirety or in part — with the content or performance expectations of the standards.**

Massachusetts made a wise decision in developing its own assessments rather than purchasing "off the shelf" tests. The vast majority of the items test content specified in the standards. When reviewers used the state-provided map, seven items were identified as not aligned to the standards, meaning that the mathematics of the item is peripheral to the content of the standard. However, reviewers were able to identify a more appropriate match for three of these seven items to other

expectations found in the state's standards. Consequently, in the final analysis, over 90 percent of MCAS test questions are aligned to content in the standards. This finding compares favorably with assessments in other states Achieve has reviewed, where in some cases a substantial number of items are not found in the standards at all or are very weak matches to the standards.

Reviewers also matched the performance required by the item to the performance described in the standard. As was the case with content, a large proportion of test questions (86 percent, if the original map is assumed) measure skills specified in the standards. If the reviewer-revised map is used, alignment improves to over 90 percent. Only one item received a score indicating that the related standard is stated in such general terms that reviewers could not be certain of direct alignment. This reflects highly on the Massachusetts standards, which tend to be written in clear and measurable language that specifies the precise skills that students are to learn.

3. **The grade 10 MCAS test in mathematics is a fair and appropriately challenging test for high school. Students who perform below the "needs improvement" level (currently a score of 219 or lower) likely have a minimal level of skill in mathematics and will need intensive instructional support to achieve the minimum standards.**

The level of challenge for the grade 10 assessment, as a whole, is appropriate when compared with the 8th- and 10th-grade standards to which the items are mapped. The test is more challenging than most other mathematics high school tests Achieve has reviewed, to be sure, but that is because many states have high school tests in mathematics that tend to be pitched at the grade 8 level.

Thirty-one percent of the 2001 grade 10 MCAS common items, or a total of 13 questions, were specifically designed to assess 8th-grade standards. Including such a large proportion of items that assess grade 8 standards has an impact on such alignment criteria as level of demand, balance and range. It also has important policy implications for Massachusetts, since a student who can correctly answer all of the 11 multiple-choice items and the single short-answer item, and get at least some points on the single open-response item, will receive from 13 to 16 points on the assessment. Only 21 points were required in 2000 for a student to "pass" the MCAS mathematics assessment by achieving the 220 "needs improvement" standard. This issue — and the fact that many of the less cognitively demanding aspects of the 10th-grade standards are assessed, rather than those aspects that are more challenging — contributes to Achieve's view that students who do not achieve the 220 score are minimally competent in mathematics.

The item set mapped to Patterns, Relations and Algebra tends to assess the more cognitively demanding aspects of the related standards, while still being at an appropriate level of challenge overall for 10th graders. It contains a rich mix of item types — five multiple-choice items, two short-answer items and three open-response items — and presents students with a higher level of challenge than do the item sets for the other strands. The mix of item types, with a relatively heavy emphasis on open-ended items, affords students the opportunity to demonstrate their abilities to not only model and solve equations and inequalities, but to also apply those skills to the solution of everyday problems.

The level of challenge of the items mapped to Geometry is appropriate for all 10th-graders, even though it is less demanding overall than the Patterns, Relations and Algebra questions. Several standards require students to use specific geometric knowledge and properties to solve problems, and the assessment follows through, requiring students to not only solve problems in a purely mathematical context, but also to solve everyday problems. The reviewers also deemed the level of challenge of the four items assessing Measurement concepts as appropriate, though a bit unbalanced. The four items all require students to apply formulas to calculate area or circumference, and three of the problems are presented in the context of a real world problem.

However, the items measuring Number Sense and Operations tend to assess the less cognitively demanding aspects of the standards. They place too much emphasis on numeric and symbolic manipulation, at the expense of application of number sense and computational skills to solve problems. In Achieve's view, a number of these items were designed to test Number Sense concepts, but instead more directly measure standards in the Algebra strand. This has the effect of weighting the test overall more heavily toward algebra and omitting more advanced number concepts that could be assessed.

Reviewers also expressed concern with the overall challenge of the set of items mapped to Data Analysis, Statistics and Probability as compared with the expectations laid out in the grade 10 standards. Their concern was based on the fact that so many of the items — seven of eight — assessed grade 8 standards at the expense of 10th-grade standards.

4. **The grade 10 MCAS 2001 is a technically well-crafted test; this is no small feat, considering that the state department of education works with teachers to create an entirely new assessment each year.**

Reviewers were pleased to find very few instances of such flaws as "trick" items, misleading graphics, multiple or no correct answers, or ambiguous directions. In addition, the distracters used in multiple-choice questions were plausible, meaning that students generally had to know the content to arrive at the correct answer and not get it by default because alternative choices were too far off the mark. Only two (4 percent) of the items caused reviewers concern. This contrasts with some other states, where as many as 25 percent of items on an assessment were problematic.

5. **The MCAS test is a solid measure of students' ability to reason, solve problems and connect various mathematical ideas. The assessment appropriately uses various formats — multiple choice, short answer and open response — to assess students' procedural, conceptual and problem solving skills.**

Importantly, we found that the grade 10 test does a good job of assessing students' conceptual knowledge, even though conceptual understanding is not as strongly developed in the state's standards as are procedural and problem solving skills. This is due in part to the state's judicious use of test formats. The multiple-choice items on the grade 10 assessment are generally strong:

They address important mathematical content, assess students' abilities to recall and apply simple algorithms, and test their skills in reasoning and solving more sophisticated problems.

One multiple-choice item, for example, requires students to demonstrate a more thorough understanding of the concept of "mean" than mere application of the formula for finding the "average." This particular item requires students to have a conceptual understanding of what the "mean" is so they can determine the impact of a change in one or more values of the data set. Massachusetts also effectively uses multiple-choice items to assess students' abilities to solve problems that require them to integrate different areas of mathematics. Item 9, for example, requires students to integrate their understanding of geometry (specifically, area of a circle) with probability to correctly solve a real-world problem.

Massachusetts should be commended on its use of a significant number of short-answer and open-response items on a large-scale, on-demand assessment. Out of a total of 51 test items, four are short answer, and six are open response. This amounts to 20 percent of the assessment and 40 percent of the score points. Including open-response items, rather than relying solely on a multiple-choice format, has a major effect on the way teachers teach and construct classroom assessments and the way in which students prepare for tests. Open-response items afford students the opportunity to demonstrate a deeper and more sophisticated understanding of content. Items of this nature provide opportunities to assess higher-level skills — such as students' abilities to construct graphs, write equations and explain their answers — which is usually impossible with multiple-choice items. Such items also provide a deeper venue to assess students' abilities to integrate multiple aspects of mathematics. In addition, the item-specific rubrics for all open-ended items assure that students are measured against meaningful criteria.

As a case in point, one of the open-response items on the assessment requires students to use data on alternative billing plans for a cellular phone to calculate a hypothetical monthly bill, derive an equation that generalizes how to calculate such a bill, construct a graph showing the monthly bills for a range of minutes, use the graph and/or equation to find the number of minutes of phone use for which two plans have the same cost, and explain how they came up with their answer.

Similarly, another open-response item — item 20 — requires students to use tabular data on the monthly income and expenses of a software company to construct graphs, estimate and predict income for a given month in the future, determine in which future month profit will exceed a given value, and explain how they determined their estimates.

Short-answer items can assess some essential content effectively and do it more efficiently than can open-response items. They also can tap a different level of skill than multiple-choice items. MCAS uses, for example, one short-answer item that requires students to write an equation that can be used to figure out how many passengers are on a bus before the first stop. While such content could be assessed with a multiple-choice format, using a short-answer item forces students to derive the equation, rather than select it from a list of four choices.

Areas for Improvement

Achieve's reviewers identified the following areas where future editions of the grade 10 MCAS 2001 mathematics test could be strengthened:

1. **While the state's "test blueprint" that matches individual test items to specific standards is generally good, a more accurate and detailed map might be helpful in future assessment development efforts.**

The Massachusetts Department of Education provided Achieve with a "test blueprint" that maps each of the 42 common items to the single learning standard in the 2000 framework that it is intended to assess. Achieve reviewers were provided with this map and asked to score items based on it. While Achieve reviewers generally agreed with the mapping of items to learning standards provided by the state, they also agreed that some items more closely measure learning standards other than the ones the state mapped them to. In particular, a number of test questions mapped by the state to the Number Sense strand are actually, in Achieve's view, more strongly aligned to Algebra standards; this results in a disproportionate number of items assessing the Algebra strand. Reviewers' concern is not about the items' quality or utility, but rather that the test tends to emphasize algebra concepts more than number concepts and is thus out of sync with the item distributions published in the *Overview of the MCAS 2001 Tests*.

The *Overview* document also explains that "some items incorporate standards identified for preceding grade levels … (e.g., grade 10 students may be tested on learning standards identified in the framework from pre-Kindergarten through grade 10)." This is a worthwhile strategy, particularly when states assess knowledge and skills at checkpoint grades, such as 4, 6, 8 and 10. It sends the message that the discipline builds on the understandings and skills of previous years and that students are expected to carry over that knowledge from year to year. However, Achieve's reviewers were concerned that such a large proportion of the 10^{th}-grade MCAS items measure middle school standards. The detailed map provided to Achieve of 10^{th}-grade test items to the state standards shows that 29 common items (69 percent) map to grade 10 learning standards, and the remaining 13 common items (31 percent) map to 8^{th}-grade standards.

This issue can be partly addressed in future rounds of testing. Massachusetts currently identifies only "primary" mappings of items to standards, i.e., one item measures only one standard. In Achieve's view, many test questions on MCAS actually assess more than one concept, often from different standards, so the state may wish to adopt the practice used in some other states and identify both primary and "secondary" mappings of items to standards. In practice, this means that an item might be noted, for example, as assessing both a 10^{th}-grade geometry standard and an 8^{th}-grade algebra standard. Such connections across strands are already evident in the test questions and should be explicitly communicated to educators and families in the *Overview* and other public assessment documents.

2. The balance of the grade 10 MCAS 2001 is somewhat uneven; this tends to depress the test's level of rigor.

While the item sets representing Number Sense and Operations, as well as Patterns, Relations and Algebra, tend to fairly represent the balance of content knowledge and skills expressed in the standards, the item sets for Geometry, Measurement, and Data Analysis, Statistics and Probability are less balanced.

The finding that four of nine items mapped to Number Sense and Operations did not actually assess the content expressed in the standards somewhat diminished the balance of the set and de-emphasized the more advanced number concepts that should be part of the high school curriculum for all students.

Reviewers commented that the balance of the set of items mapped to Patterns, Relations and Algebra would have been improved if standard 10.P.2 and the part of 10.P.3 that requires students to "Add, subtract and multiply polynomials" (allowable for testing in 2001) had been assessed. They observed that both standards contain key concepts, critical to the development of a strong foundation in algebra. Standard 10.P.3, for example, includes such concepts as using graphical and algebraic representations to determine slope and x- and y-intercepts *and* determining a linear equation from a graph or geometric description.

The Geometry item set is less demanding than the standards in many ways, as a number of key geometric concepts to which all students should be held accountable are omitted. For example, standard 10.G.4 requires students to "apply congruence and similarity correspondences ... and properties of the figures to find missing parts of geometric figures ..." Other standards — including 10.G.2, 10.G.3, 10.G.4, 10.G.8, and 10.G.10 — are also excluded from the assessment.

Reviewers also found that the balance of the set of items mapped to measurement needs improvement; some key measurement concepts identified in the grade 10 standards are not assessed by this test. Three of the four measurement items require area calculations (involving triangles, squares, rectangles and circles), and two of the four require students to be able to calculate circumference (formulas are provided on the Mathematics Reference Sheet). Students are not required to find the surface area or volume of three-dimensional shapes such as prisms, pyramids, cylinders and cones, as specified in standard 10.M.2. Likewise, the test does not require students to "relate changes in the measurement of one attribute of an object to changes in other attributes, e.g., how changing the radius or height of a cylinder affects its surface area or volume" (standard 10.M.3).

The mix of items assessing the strand on Data Analysis, Statistics and Probability is good. It includes items requiring students to address all three areas (Data Analysis, Statistics and Probability), and the items incorporate a variety of types of graphical and tabular data displays. That said, reviewers pointed out that while this strand contains three standards at grade 10, only one of the three 10th-grade standards is assessed. Instead, seven of the eight items mapped to this

strand actually assess grade 8 standards. The reviewers were not troubled by the concept of cumulative progress — assessing important skills and concepts from earlier grades — which they found to be reasonable. Their concern was rather that such a large proportion of items in this one strand were mapped to grade 8 standards, and that the grade 8 standards, in this case, tend to be more foundational and descriptive than those in grade 10. As a result, important content in the grade 10 standards ended up not being assessed. For example, students are not required to "approximate a line of best fit for a given set of data" (standard 10.D.2) or to "describe and explain how the relative size of a sample and the population affect the validity of predictions from a set of data" (standard 10.D.3).

ENGLISH LANGUAGE ARTS

Although the Massachusetts English language arts (ELA) Curriculum Framework (adopted November 2000) has now been revised, the MCAS ELA 2001 tests are still based on the 1997 frameworks. 2002 test forms will be based on the revised frameworks. These learning standards were developed in collaboration with teachers, school and district administrators, reading and writing specialists, college faculty, and parents.

The language and literature portion of MCAS assesses learning standards number 4 through 17 of the framework's language and literature strands by requiring students to read a selection of literary and nonliterary passages and then respond to a set of multiple-choice and open-response items based on each passage. Approximately 50 percent of the passages are by authors listed in Appendices A and B of the 1997 framework. The composition portion of the test assesses students' skill at literary analysis, requiring them to use their knowledge of literary elements, themes and structures to analyze an excerpt from a literary text.

Strengths of the Grade 10 Assessment

Achieve's reviewers identified the following strengths in the grade 10 MCAS 2001 English language arts test:

1. **Massachusetts is committed to sharing assessment information with all stakeholders; what is tested is not a state secret.**

In its February 2001 document, *Overview of the MCAS 2001 Tests*, the state presented the distribution of item types and points for the three English language arts strands (Language, Reading and Writing). Overall, the 2001 English test adheres closely to those descriptions. Moreover, the percent of total points allocated to each item type is fair and appropriate across the test. This public presentation of the test elements is not common to all state assessment systems and is an important part of the mutual accountability effort displayed by Massachusetts.

Reviewers were impressed with the clarity and amount of information provided by the state on its assessments and commend the state for providing such easy access to information. The detailed feedback provided by MCAS reports at the student, school and district levels helps ensure that the testing is beneficial for individual diagnosis and for curriculum and instructional planning. Massachusetts is also to be commended for the release of its common test items so promptly after test administration. The release of these items helps ensure public credibility and provides concrete models for teachers to use as they design and develop instructional activities. Such a high degree of public accessibility to information about a state assessment system is rare, yet essential to sustaining public support for higher standards.

2. **The match between the assessment items on the grade 10 test and the Massachusetts standards is strong; the test only addresses content or skills included in the standards.**

All the test items align with the standards to a greater or lesser degree, and everything on the test can be found in the state standards. In terms of the content match between each test item and its related standard, of the 34 items assessing reading, 28 were strongly aligned or partially aligned in that they captured only part of the standards' intent. The remaining six items were partially aligned because the standards to which they were mapped are rather imprecise.

The grade 10 MCAS in English language arts received the highest scores in reading in performance centrality — the match between the performance required by an item compared to its related standard — that Achieve has seen in looking at more than 25 large-scale language arts assessments to date. Of the 34 items, 30 were scored as strongly aligned, indicating that the performance described in the standard is very consistent with that required by the item. When the standard required identification, the items required identification. When the standard required analysis, the items did as well.

3. **The grade 10 MCAS 2001 test in English language arts is a rigorous, fair assessment of the Massachusetts standards and is set at a level appropriate for high school students.**

With a few minor exceptions, the items in this assessment are very strong, requiring a level of literacy appropriate for 10th grade. This test presents an admirable set of expectations for the high school student and is, in this manner, faithful to the intent of the rigorous standards upon which it was based. Students who score well on this assessment can be regarded as having masterful control of the elements described in the standards, and students who do not score well on this test need strong instructional support in order to master the basic skills required by the English language arts standards.

Moreover, unlike numerous state assessments reviewed by Achieve, the grade 10 test is a reasonable measure of the level of proficiency in English language arts reviewers would expect of high school students. Whereas some state assessments tend to over address low-level reading skill — even at the high school level — this assessment devotes the majority of its items to the assessment of higher-order comprehension skills, a demand that resonates with the expectations described in the Massachusetts Curriculum Framework.

Level of demand (1, 2, 3 or 4) is an alignment criterion that reviewers use to gauge the type and kind of thinking required by a student in answering an item. Level 1 items tend to call for basic identification or literal comprehension. Level 2 items generally demand the student to draw inferences. Level 3 items demand more in the way of reasoning and interpretation. And level 4 items usually require complex and extended analysis. (The level 4 category of evaluation and sophisticated analysis is typically reached only in an item that is an extended response, requiring some research and use of multiple references — a level not probable in an on-demand, large-scale situation.)

The levels of demand for the reading and vocabulary items ranged from 10 items scored at the literal/recall level (1) to 26 items scored at the inference and simple analysis level (2), to all four open-response items scored at the interpretation level (3). This range is commendable for a 10th-grade test. High school students should be expected to be able to go beyond basic comprehension, and this test asks them to do so. In comparison, other states have tended to pitch the majority of reading and vocabulary items at the level 1 category.

The 2001 grade 10 MCAS offered the students one writing prompt: "A frequent theme in literature is the conflict between the individual and society. From a work of literature you have read in or out of school, select a character who struggles with society. In a well-developed composition, identify the character and explain why this character's conflict with society is important." To determine whether the writing prompt was set at an appropriate level of challenge — neither too hard nor too easy for a 10th-grade student — the reviewers studied the anchor papers provided for this prompt. Anchor papers are examples of the various score ranges — in this case from 1 to 6 for topic development and 1 to 4 for use of conventions. It is in the selection of the anchor papers, those used by the scorers to determine score ranges, that the standards of performance are set.

In looking at these anchors, the questions for the reviewers were "How hard is it to get a passing score?" and "What is the level of writing skill that is required?" Reviewers determined that although the prompt presented a real challenge to the student, the score points established were very reasonable. In essence, a very high level of response was necessary to score a 6, but only a moderate response was necessary to score a "passing grade" of 4. The scoring is therefore set at an appropriate level of challenge for 10th-grade students and is in line with the expectations described in the standards. Students who earned 5 or 6 points for topic development are strong writers and literary analysts, while students who scored less than 4 points were minimally competent in these areas. Achieve's reviewers made the same judgment with respect to the scoring of conventions on the writing assessment.

4. The 2001 grade 10 MCAS is a well-crafted test. No items have technical difficulties in the way they are constructed, though the state may wish to revisit the choice of writing topic in future editions of the MCAS.

All of the items in both the language and the reading strands are soundly constructed. No items are trick questions, have implausible distracters or address elements of the texts that are so insignificant as not to be "fair game." Unfortunately, such unfair sources of challenge too often are used in other assessments that Achieve has reviewed. This test's language and literature strands were judged to be free from any elements that may cause a student's performance to be due merely to test-taking skills rather than to accomplishment in the subject matter.

Regarding writing, Achieve's reviewers have some concerns. The two basic criteria for the assessment of the writing sample — focus and development and use of standard English conventions — are detailed in the student test booklet, thereby giving the student clear direction

in his writing. The essays are scored against a generic rubric, which is a public document. This is a strong element of the assessment in that it makes the demands of the test public.

However, Massachusetts' particular choice of writing topic prompted some debate among the reviewers. The fact that students are allowed to choose a character from anything they had read, either in or out of school, was regarded as a plus. Such choice makes the topic accessible to students coming from a variety of curriculum and independent reading experiences. As is always the case, however, choice is a mixed blessing for the student test taker. Not all texts present a character who struggles with society, so a major part of the difficulty inherent in this topic is the selection of an appropriate text to discuss. As is evident in the anchor papers for this prompt, student selection of text often determined how successful the writing would be. A wise selection of a text increased the likelihood that the student's essay would be successful, whereas a poor selection diminished the essay's chances of being successful.

The explication of theme also presents a source of challenge in this prompt. Although theme is sometimes used in the same sense as motif to signify recurring concepts in literature, such as the conflict between the individual and society, the term more generally refers to the argument or overall idea expressed by a specific literary work, whether implied or explicitly stated.[3] These two different definitions of theme, the universal and the text specific, require different levels of abstraction when applied to a text.

It is the text-specific definition that is tested when students are asked in a multiple-choice item to determine that the first sentence of Loren Eiseley's "The Angry Winter" introduces the theme of the essay. ("The time comes when creatures whose destinies have crossed somewhere in the remote past are forced to appraise each other as though they were total strangers.") In contrast, the writing prompt tests students' ability to apply the highly abstract concept of universal themes. It is true that standard 11 requires students to "apply knowledge of the concept that theme or meaning of a selection may involve several ideas and then analyze and compare works that express a universal theme, providing evidence to support their ideas." Nonetheless, although this expectation is clearly described in the standard, the degree to which *specific* instruction in universal themes had been provided for the student may well have helped determine how successful s/he was in writing an essay on this prompt. In summary, a high score on this writing assessment does not only suggest a strong writer; it suggests a strong writer of literary analysis. The composition score on this test reflects as much, if not more, literary analytic skill as writing skill.

5. The grade 10 MCAS test in English language arts uses multiple kinds of item formats to good advantage.

The grade 10 English language arts test makes good use of multiple formats — multiple choice, open response and a writing sample — that are appropriately employed for the skills being tested. The multiple-choice items are unusually strong, being free of confusing formats and addressing important aspects of the reading passages. Additionally, the open-response items were used well in

[3] *The University of Victoria Writer's Guide*, The Department of English, University of Victoria, 1995.

this test. As noted previously, all four open-response items were judged by reviewers to be level 3 interpretation items. Students may need to apply knowledge of concepts to the text in order to determine meaning, explain, generalize or connect ideas and support their thinking. The items require a deeper level of analysis than multiple-choice items are capable of and the item-specific rubrics assure that the student responses are measured against meaningful criteria. Moreover, the rubrics are clear, and their wide publication should prove helpful for students who are preparing for this kind of assessment.

6. The quality of the reading passages on the grade 10 test is quite high.

Reviewers were impressed by the quality of the reading passages. The texts are appropriately complex for the grade level, and their sophistication made it possible to generate effective multiple-choice and open-response items. Another plus is that at least 50 percent of the passages were chosen from the list of authors included in the appendix in the curriculum framework and also represented a range of time periods. This contrasts with other states Achieve has reviewed, where the high school test's reading passages too often were of lesser quality. Massachusetts has run numerous reading-level scales on all of the passages, and they all fall well within a high school range.

In view of the fact that the state added a standard focused on drama in the 2000 standards revision, Massachusetts might wish to direct attention to Shakespeare's works — a notable part of every high school English curriculum — by using an excerpt from a play, instead of a sonnet.

Areas for Improvement

Achieve's reviewers identified the following areas where future editions of the grade 10 MCAS 2001 English language arts test could be strengthened:

1. The state should improve the balance of fiction and informational text included on the test.

All students are entitled to rich exposure to both literature and informational text. In the initial stages of the standards movement, many leaders, particularly those in business and industry, made the case that students should be proficient in reading and writing a broader range of text than traditionally studied in American classrooms — the kinds of text they would confront as citizens, employees and consumers. Nearly all state English language arts standards, including those of Massachusetts, expect schools to sample the spectrum of expository, functional, historical and public documents (e.g., speeches, historical documents, informational articles, journal articles from other academic disciplines, instructional manuals, editorials, political essays) in addition to the study of literature that includes fiction (short stories, novels, drama, poetry) and nonfiction (biography, autobiography, expository essay writing).

Each of these kinds of texts makes very different demands on the reader, writer, speaker or listener, and state assessments should include reading passages and test items that probe the various demands. In response, most states have chosen to include standards that explicitly call for students to be able to read and write various kinds of informational text and have also modified their assessments to include reading passages and test items that assess their ability to do so.

The MCAS grade 10 assessment in language arts is heavily weighted toward the interpretation of literary elements, however, whether they are in fiction or nonfiction texts. Such weighting is not necessarily a negative, in that the standards themselves are so weighted. What is slighted in this assessment, though, is a balanced attention to the nature and qualities of informational texts.

In the 2000 revision of the curriculum frameworks, a clear demarcation is described between imaginative/literary texts and informational/expository texts:

> **Standard 8: Students will identify the basic facts and essential ideas in a text and use them as the basis for interpretation.**
>
> Grades 9–10:
>
> For imaginative/literary texts:
>
> 8.29 Identify and analyze patterns of imagery or symbolism.
>
> 8.30 Identify and interpret themes, and give supporting evidence from a text
>
> For informational/expository texts:
>
> 8.31 Analyze the logic and use of evidence in an author's argument.

In developing the 2002 tests, Massachusetts should ensure that this additional emphasis on informational texts is reflected in both the passages selected and the kinds of questions asked, with an increased attention to argumentation and structure in expository materials. All of the passages on the current test are either fictional texts (a novel excerpt and two poems) or nonfiction (two essays and one autobiography excerpt). Although technically classified as nonfiction, the essay, "The Angry Winter," and the excerpt from *The Autobiography of an Ex-Colored Man* are both narrative in structure and written in a language rich in symbol and metaphor; both excerpts are, thus, highly literary in tone and style. In Achieve's view, no informational or expository text is represented is this assessment.

2. The state may wish to consider mapping items to more than one standard, where appropriate.

In the reading strand, reviewers noted many items that address multiple standards due to the over-lapping content of the standards and the relatively sophisticated nature of the test questions. For example, only one item was mapped to standard 11, "Students will identify, analyze and apply knowledge of theme in literature and provide evidence from the text to support their understanding." By devoting a separate standard to theme, the state highlights the importance of this literary element and should, therefore, assess it with more than one item, especially since understanding the theme of a text really means understanding the meaning of the passage — the

acid test of comprehension. And in reality, Achieve's reviewers found that other items on the 2001 test do assess theme, but are mapped to other standards. Consider these four open-response items:

> **Item 8:** Mark Twain said, "Make your vocation your vacation." Explain how this quotation relates to this article. Use specific evidence from the article to support your answer.
>
> **Item 15:** "Sonnet 116" does not have a title linked to the text; rather its title distinguishes it from Shakespeare's other sonnets.
>
> - What title would you give to "Sonnet 116"?
> - Provide evidence from the poem to support your answer.
>
> **Item 42:** Explain the significance of the statement in lines 82 and 83, "It was he who was civilized now," as it applies to both the man and the dog. Use specific evidence from the essay to support your answer.
>
> **Composition Topic:** A frequent theme in literature is the conflict between the individual and society. From a work of literature you have read in or out of school, select a character who struggles with society. In a well-developed composition, identify the character and explain why this character's conflict with society is important.

So, although the state map of the items makes it appear that standard 11 is under-represented with just one item, theme *is* adequately assessed. The present mapping strategy does not reveal this appropriately strong emphasis. The standard devoted to theme is a case in point. If an item addresses the theme of a nonfiction text, the state maps it to standard 13, the nonfiction standard. If an item addresses theme in poetry, the state maps to standard 14, the poetry standard. Yet in each case, the item actually addresses two standards — the theme standard and either the nonfiction or poetry standard. If these items were double-mapped to both standards, the message about balance in the test might be more accurate. Similar situations occur with standards 10 and 15.

In summary, even though the mapping of the test items to the general standard is appropriate for this assessment, the standards themselves overlap to the extent that distinguishing which item relates to which standard is often "guesswork," and some items may be legitimately mapped to more than one standard.

3. Massachusetts should consider adopting an item-specific rubric to score the writing sample on the grade 10 test.

Although item-specific rubrics are used to score the open-response items on both the mathematics and the language arts tests, the state chose to use a generic rubric to score the writing sample. The generic rubric addresses four elements of topic and idea development (idea development, organization, details and language/style) and three elements of the use of conventions (sentence structure, grammar and usage, and mechanics). Scorers are encouraged to provide analytic annotations for each element, noting either a commendation or a need. The state should be

commended for trying to provide detailed feedback to students and schools by means of analytic annotations. The development of an item-specific rubric for the writing prompt would complement this effort by helping students who are learning to write an effective literary analysis and more clearly communicating the expectations for a composition of this type to the public.

4. While Massachusetts deserves praise for assessing writing directly, the state may wish to reconsider the effect of confining the writing prompt to a single genre.

Standard 20 requires that "Students will select and use appropriate genres, modes of reasoning and speaking styles when writing for different audiences and rhetorical purposes." The 1997 Curriculum Framework, additionally, describes five "common modes of presentation." These include exposition, narration, argumentation, exposition, description likened in various ways to the four "aims of discourse," which are informational, persuasive, expressive and literary (p. 53). A variety of writing genres, therefore, is described in these curriculum frameworks.

Reviewers applauded Massachusetts' decision to assess compositional and convention skills directly. Developing a significant piece of writing takes time, however, and it is typical in large-scale testing that only one genre of writing is required per student, even though the standards call for a variety of writing types. It is also typical that the genre assessed on state tests varies over test forms: Pennsylvania uses a matrix of topics, meaning that different groups of students respond to a variety of prompts so that in any one year a student cannot predict the type of prompt s/he will confront; New York assesses a variety of writing types over the years; Oregon provides students with a choice of forms and topics each year. Massachusetts has instead chosen to identify a particular writing type that will be assessed each year for each grade tested: narrative at grade 4, persuasive at grade 8 and literary analysis at grade 10. The decision to designate only one type of writing for state assessment has both positive and negative implications for instruction.

On the plus side, students and teachers preparing for this 10th-grade exam know ahead of time just what the genre or mode of the prompt will be. They know that a clear and steady focus on literary analysis, at least at grade 10, will be the best preparation for the assessment. The required format will present no surprises; the requirements are public. In short, announcing the type of writing that will be assessed allows teachers to focus on that type of writing in their classrooms.

On the other hand, since testing often does drive instruction, other states have opted to vary the writing formats required on the yearly assessments. This helps ensure instruction in a variety of writing types described in the standards, so that no one genre receives undue instructional emphasis. One might argue that literary analysis is the proper genre for study in an English class — both traditionally and in practice — and is, therefore, the most important genre to be tested at the high school level. Reviewers observed that, in effect, the composition score for the 10th-grade MCAS is a determination of a student's ability to write a literary analysis. This is in contrast to states that either matrix the prompts or vary the genre from year to year, where the writing score is more of a generic assessment of a student's ability to approach any writing form.

Major Findings: Massachusetts' Mathematics Curriculum Framework

The Massachusetts Mathematics Curriculum Framework was adopted by the Board of Education in July 2000, replacing the 1996 Massachusetts Mathematics Curriculum Framework known as *Achieving Mathematical Power*. The 2000 framework presents revised state guidelines for learning, teaching and assessing mathematics, and reflects the hard work and dedication of public school teachers and administrators, mathematics education professors, university mathematicians and community members from across the commonwealth. Such a periodic review and revision of standards documents is required in Massachusetts by the Education Reform Act of 1993.

The 2000 framework is a significant improvement over the 1996 framework in several regards. First, the grade structure used to organize the two documents is tighter and more helpful for curriculum planning. The 2000 framework organizes student standards into bands consisting of two grades: Pre-K–K, 1–2, 3–4, 5–6, 7–8, 9–10 and 11–12, whereas the 1996 framework used broader bands — Pre-K–4, 5–8, 9–10 and 11–12. Secondly, the 2000 framework restructured the strands to more closely represent major domains of mathematics, adding an additional strand to separate measurement from geometry. Third, while neither the 1996 nor the 2000 document includes a separate strand for mathematical processes, both attempt to embed important mathematical processes into the content strands. The 2000 framework, however, describes five mathematical competencies: problem solving, communicating, reasoning and proof, making connections, and representations, whereas the 1996 framework focused on four such competencies: problem solving, communicating, reasoning and connecting. This is consistent with the recommendations made by the National Council of Teachers of Mathematics (NCTM), when each of the respective framework documents was drafted.

A number of other improvements have been made in the 2000 framework, with some being design changes and others being more substantive. For example, the revised framework contains a more streamlined introductory section and omits the vignettes used in the 1996 framework to depict what the standards might actually look like when implemented in the classroom. Both the 1996 and 2000 frameworks contain concrete examples that help translate the standards into practice, but the nature of the examples is different. The 1996 framework contains Examples of Student Learning that tend to be brief narrative descriptions of student activities or problems, whereas the 2000 framework provides Selected Problems or Classroom Activities that are more focused and are linked to specific standards.

The 1996 and 2000 frameworks also reflect differing philosophies toward the use of technology in the mathematics classroom. The 2000 framework is much more clear and direct about expectations relative to technology when it states: "... calculators should not be used as a replacement for basic understanding and skills ... Elementary students should learn how to perform thoroughly the basic arithmetic operations independent of the use of a calculator ... Although the use of a graphing calculator can help middle and secondary students to visualize properties of functions and their graphs, graphing calculators should be used to enhance their understanding and skills rather than

replace them." This philosophy is consistent with research showing that students can learn mathematics more deeply with the use of appropriate technology.

As noted previously, Achieve asked a cadre of five mathematics educators and mathematicians to review and benchmark Massachusetts' mathematics standards. The reviewers were selected because of their deep knowledge and experience with mathematics, mathematics education and Pre-K–12 mathematics standards — and because they represent a range of perspectives relative to mathematics that, when taken collectively, provide the state with a well-rounded, well-informed perspective on the standards. The charge to the reviewers was to critique the standards using a set of guiding questions drafted by Achieve and compare them to other identified exemplary mathematics standards documents: those of Japan, Arizona and Achieve's Mathematics Achievement Partnership (MAP), of which Massachusetts is a founding partner state.

It is important to place Japan and Achieve's MAP standards, limited to the middle grades, in the larger context of mathematics education reform across the states. Achieve's MAP standards look to the future; no state has yet articulated standards of equal rigor, and no state is yet ready to implement and hold students accountable for what is admittedly a high level of achievement. For this very reason, Achieve's MAP partnership with states like Massachusetts entails more than formulating a set of exemplary end-of-grade 8 standards; Achieve will also generate supportive professional development and identify aligned curriculum materials. Achieve's benchmarking process is designed to compare a state's standards to some of the best in the world. It is therefore unrealistic to expect a state's standards, its teachers and its students to align with these external benchmarks overnight. The message has been and remains one of "continuous improvement." Consequently, Achieve's analysis highlights areas in which Massachusetts does particularly well and areas where additional work may be warranted as the quality of both teacher and student preparation improves.

Massachusetts deserves high praise for its diligent efforts to improve its 1996 set of mathematics standards. One reviewer went so far as to say that, all in all, the Massachusetts mathematics standards are now one of the best in the country. That said, Achieve's standards are very high, and our recommendations for improvement are made in the spirit of how Massachusetts can, over time, take its mathematics standards to still another level, making them "world class."

Strengths of the 2000 Mathematics Standards

1. The standards are clear, measurable and jargon-free.

Massachusetts' standards use clear, specific and concrete language to define for educators and the public what students should know and be able to do. In terms of clarity and specificity, these standards are comparable to those of Arizona and superior to Japan's. There are some references that use specific mathematical language that may not be familiar to the general public (e.g., *Pythagorean Theorem* or *transversals of coplanar lines*), but these are appropriately used and would be very difficult to avoid.

The state also does an excellent job of not only defining what content students should know at specific grade spans but also what type of measurable skill (e.g., identifying, estimating, predicting, graphing) students need to be able to demonstrate. Clear and measurable standards such as these provide a solid foundation for teachers as they develop activities and assessments for their students and for test developers who work on large-scale assessments such as MCAS.

2. The guiding philosophy, guiding principles and strand overviews are clear and explicit.

If the introductory sections of a standards document are to be read and be useful to practitioners, brevity and clarity are key. Massachusetts has been successful in condensing its introduction, guiding philosophy, guiding principles and strand overviews into less than 15 pages. This is a reasonable length, while still being sufficient to lay the context and set the parameters for the document.

The revised introductory sections are substantive and clear. The Guiding Philosophy consists of five well-stated processes (problem solving, communicating, reasoning and proof, making connections, and representations) that appear as standards in other standards documents. They are positive additions and are included in the lead-in statement to every strand in the document. The Guiding Principles provide a helpful framework to describe the environment in which mathematics learning should occur. They address the areas of learning, equity, teaching, technology and assessment, and send the message that such issues must be attended to if students' opportunities to learn mathematics are to be optimized. Similarly, the Strand Overviews are compelling and provide a vision of the trajectory of learning within a strand that will be expected of students as they progress through the grades.

3. The inclusion of selected problems or classroom activities and a checklist of exploratory concepts and skills is an excellent strategy for making the standards concrete and clear to teachers and parents.

Relatively few states effectively use sample problems and activities to clarify the meaning of their standards for teachers and the public, and Massachusetts is to be commended for employing this strategy. Since the sample problems included in the 2000 framework are specifically linked to one or more standards, they are particularly effective in anchoring the level of expectation of the standards. As one reviewer noted, "Only by providing examples of sample problems can one determine the true meaning of the standard."

The Exploratory Concepts and Skills should also be helpful to teachers. They preview the direction in which learning will be going as students progress to the next level and call teachers' attention to ways to broaden and enrich instruction and make it more meaningful. They have the potential to be particularly useful as a guide to accelerated teaching and learning for teachers working with students ready to move beyond grade level expectations. One reviewer suggested that Massachusetts consider adding pedagogical commentary similar to that in Japan's standards to further enrich the Exploratory Concepts and Skills.

4. The standards are comprehensive, and the amount of content is reasonable.

The 2000 Standards include the most important content for grades Pre-K–12, and the amount of material presented in the learning standards is reasonable. Conversely, most of the content in the Massachusetts standards also appears in the benchmark documents; that is, there are not a lot of "extraneous" expectations in the Massachusetts document relative to the benchmark documents. Differences that do occur tend to be not so much in what content is included but rather in the ways that content is encountered. For example, while the Massachusetts standards include a mix of procedural, conceptual and problem solving expectations, they are more heavily weighted toward the procedural than are the benchmark documents.

5. The standards for grades Pre-K–6 generally do a good job of preparing students for more in-depth study of mathematics in the later grades.

The standards are particularly strong in the areas of computational fluency and the application of algorithms, aspects of mathematics that are important if students are to experience success in later grades. In referencing Massachusetts' Pre-K–6 standards, one reviewer noted that the subject matter is quite ambitious, with Massachusetts being well ahead of Arizona on many topics in these grades.

6. The standards generally demonstrate a clear progression of knowledge and skills from grade band to grade band.

Massachusetts clearly expects more from students as they progress through the elementary and middle grades. The numbers and geometric objects encountered increase in quantity and complexity, and number operations deal with expanding number sets and systems as students move through the grades. Number theory ideas, sequences and the use of variables all build from foundational ideas to more demanding ones as the standards are articulated for later grades. The skills demanded of students also increase in complexity. For example, in the area of geometry, students at the elementary grades are expected to name, identify, sort, draw, recognize, compare and describe. By the time they are middle school students, they are also expected to analyze, predict, explain, formulate and test, and apply geometric concepts to the solution of problems.

7. The division of the standards into two-year grade bands is effective, though there is some confusion about the use of two sets of standards in high schools.

The two-year spans strike a reasonable balance between the need to provide flexibility for local school districts in sequencing instruction and the needs to make expectations explicit and monitor progress regularly through the state's assessment system. It was not clear to the reviewers, however, that the alternative standards presented for the high school years are used most effectively (i.e., standards are presented both by course and by grade span). In particular, the course standards include content that is more advanced or goes into more depth than the standards for grades 9–10 and 11–12; Achieve's reviewers felt that a good deal of this additional content

should be expected of all students. The state may wish to clarify its intent in developing what amounts to two sets of standards: Is the purpose to provide flexibility in curriculum choice to schools and families or, indeed, to articulate two different sets of standards?

Recommendations for Improvement

1. Massachusetts should strengthen its standards by emphasizing the critical role that mathematical reasoning plays in connecting the domains of mathematics to each other.

Although Achieve's experts agreed that Massachusetts lays a strong foundation in procedural understanding in the early grades, they believe that the standards should be enhanced in the area of conceptual understanding. Japan (and China, according to Liping Ma, whose work is used as a reference in the 2000 framework) spends much time in the early grades establishing and building on the idea of place value and what it means for arithmetic operations. This lays a strong foundation for later computational work with fractions, decimals, percents and integers. Japan also pays special attention to establishing students' conceptual understanding of units of measure early on, setting the groundwork for more sophisticated measurement expectations as students progress through the grades. The Massachusetts standards, however, are lean across the grades and strands on expectations that focus on building conceptual understanding, strategic competence and adaptive reasoning. They instead tend to focus on expectations that require students to apply procedures and algorithms to do such things as read numbers, do arithmetic and measure objects.

The additional time required to build deeper conceptual understanding of arithmetic, algebra and geometry could be had by shifting the introduction of coordinate geometry and a portion of the data strand to grades 5 and 6. More emphasis on geometric and algebraic concepts is also needed at grades 5 and 6, if Massachusetts is to ensure that its middle school students are prepared for the rigors of higher-level mathematics.

Furthermore, although Massachusetts clearly and concisely addresses the importance of reasoning in its Guiding Philosophy and Guiding Principles, these topics should be infused into the standards themselves at all levels. The Arizona standards are particularly strong in this respect and might serve as a good model. For example, for grades 6–8, Arizona requires that students be able to "construct simple valid arguments using if … then statements based on geometric shapes, proportional reasoning in probability and syllogism" and to "solve problems using deductive reasoning."

In the same vein, Japan's standards for grade 8 frequently require students to "confirm" properties, implying an expectation of informal proof. The Massachusetts standards do not ask for anything like this in grades 7–8 and only hint at it in grades 9–10. Only in the separate geometry course is proof — formal or informal — listed as a standard. Interestingly enough, the 1996 framework did contain some very explicit expectations involving logical thinking and proofs in the grade 9–10 learning standards for Geometry and Spatial Sense, but these have been eliminated in the 2000 framework.

In addition, the 2000 framework, as currently constructed, does not capture the holistic nature of mathematics, its underlying structure and the relationship of one branch of study to another. Nor does it truly reflect the Guiding Philosophy's stated goal of "making connections." Reviewers offered suggestions as to how to fortify reasoning and logical thinking and transform this fragmentation. Most ideas centered on a more in-depth and strategic use of sample problems. One reviewer went further and suggested providing many problems at the end of each grade band, which are selected to illustrate the connective nature of mathematics, and which are not tied to any particular standard. These could be in addition to the Selected Problems or Classroom Activities that currently appear in the framework linked to specific standards. Such "cross-cutting" problems would also serve as models for how the mathematical processes, particularly problem solving and reasoning, can serve as a unifying thread throughout the document.

Unfortunately, some of the Selected Problems now included in the framework are not particularly representative of the breadth, focus and balance of the standards. More and better problems would be helpful to anchor the level of expectation of the standards and to make connections among mathematical ideas more explicit. The MAP expectations include extensive sample problems, which do more than show procedures or skills; rather, they demonstrate the depth of mathematical understanding and reasoning skills implied by the standards. They also typically focus on concepts that need clarification or are difficult to teach.

The intent of such illustrative problems is not to provide problems useful for assessing student performance but rather to show teachers how to help students think about the mathematics involved in a deeper way. Given that Massachusetts is a founding partner of MAP, it might make sense for Massachusetts to revisit its selection of sample problems, replacing or augmenting them to better capture the essence of the standards and support professional development.

2. **As the state's efforts in the areas of teacher development and support for struggling students intensify and take effect, Massachusetts should consider revising its middle school and high school standards to increase the level of rigor.**

Raising standards to world-class levels will take several years; Massachusetts has made many strides in this direction, but when the curriculum framework is next revised, the state should continue to inject greater depth and rigor. The level of rigor in a state's standards needs to be sufficient for students to have access to high performance jobs and/or college. The Education Trust estimates that as many as 80 percent of current students nationwide will enter two- or four-year colleges directly after high school. We realize that the expectations contained in the 2000 framework are indeed very challenging for many students. Massachusetts' standards have set a high floor, one that should be raising the level of curriculum and instruction throughout Massachusetts' schools right now. Achieve's reviewers generally agreed that the content and depth of the standards for grades 11–12 are a good foundation for success in college mathematics. Yet, they also believe that some of the key concepts now placed in these grades should be expected of all students by the end of 10th grade — regardless of their career paths — and, accordingly, that the middle school standards will need to be ramped up as well.

While a number of factors determine how rigorous a set of standards are, a principal determinant is whether the content is sufficiently challenging for the grade level. Without question, the end of grade 8 is a critical juncture in mathematics education. For students to have a genuine choice of mathematically divergent paths through their high school programs, and to not find themselves consigned to the less demanding courses by default, a state's Pre-K–8 standards must be rigorous in their conception and implementation. Algebraic and geometric knowledge and skills are gatekeepers for the two distinctly different mathematics programs common to most high schools.

Reviewers expressed concern about the rigor of content in the 2000 framework's standards for middle and high school students. Particular areas of weakness are in algebra, geometry, and probability and statistics. They generally agreed that the standards defined for grades 7–8 and grades 9–10 were insufficiently challenging compared to Japan and MAP. Massachusetts' standards for grades 7–8 in particular contain too much duplication and are too densely packed. Massachusetts might wish to shift a few of its standards to earlier grades. Currently, the standards list a number of expectations in grades 7–8 that are appropriate for grades 5–6. Japan, for example, expects students to focus on ratio and proportion at grade 6, whereas Massachusetts' students are not expected to be able to use ratio and proportion to solve problems before grades 7–8 (8.N.3). In addition, more than 30 percent of all standards for grades 7–8 focus on number sense and operations; this is too heavy a focus on number for the middle grades.

Similarly, reviewers identified a number of expectations now set for grades 9–10 that they believe are realistic for students in grades 7–8. These expectations span the various strands and include such topics as simplifying numerical expressions that may involve positive integer exponents and absolute value; understanding the geometry of lines and circles; demonstrating understanding of the relationship between various representations of a line; finding the perimeter, circumference and area of common figures such as triangles, parallelograms, trapezoids and circles; and applying congruence and similarity correspondences and properties of figures to find missing parts of geometric figures.

The fundamentally important concepts of linear and nonlinear equations deserve much greater attention in these standards. Both Japan and MAP tend to set higher expectations for middle school students for linear equations. Although Massachusetts' standards for grades 7–8 require that students be able to "set up and solve linear equations and inequalities with one or two variables, use linear equations to model and analyze problems" and "use tables and graphs to represent and compare linear growth patterns," they include no expectations for the solution of simultaneous linear equations, as do the Japanese and MAP standards. Not until grade 10 do Massachusetts' standards require students to be able to solve problems "that can be modeled using systems of linear equations or inequalities" (10.P.8/A1.P.12).

Japan and MAP also set higher expectations for students relative to nonlinear equations and functions. Japan's algebraic focus at grade 10 is on quadratic equations; by the time Japanese students reach grade 10, they are expected to have mastered linear functions. MAP standards are even more rigorous in that they expose students to nonlinear functions, and quadratic functions in

particular, in the middle grades. Massachusetts sets expectations for quadratic equations for students in grades 9–10 (e.g., 10.P.1, 10.P.5, 10.P.7), yet the approach is less ambitious than that of either Japan or MAP: The majority of Massachusetts learning standards in the patterns, relations and algebra strand for grades 9–10 reference only linear equations and inequalities (e.g., 10.P.1, 10.P.2, 10.P.6, 10.P.7, 10.P.8). To ratchet up expectations over time for student performance, Massachusetts should consider accelerating the teaching and learning of linear functions so that quadratic and nonlinear functions can get more attention in grades 9–10.

Arizona and Japan also require some knowledge of trigonometry as a core expectation for all students in grades 9–10. Arizona requires that all students be able to use the definitions of trigonometric functions to find the sine, cosine and tangent of the acute angles of a right triangle. Japanese students are expected to know about the sine, cosine and tangent functions for angles measuring from 0 to 180 degrees and to know about the law of sines and law of cosines. The 2000 framework has no trigonometry expectations prior to grades 11–12. Massachusetts may wish to reclaim the learning standard in the 1996 framework for grades 9–10 — which required students to apply trigonometry to problem situations involving right triangles — since the 2000 framework lays the groundwork by introducing students to the Pythagorean Theorem by the end of grade 8.

Another aspect of rigor is whether a set of standards exhibits a well-developed continuum of increasing complexity of knowledge and skills from kindergarten through grade 12. Reviewers generally agreed that the Massachusetts standards absolutely expect more from students as they progress through the elementary and middle grades. The one exception is the Data Analysis, Statistics and Probability strand, which contains comparatively weak standards across the grades. For example, during the elementary grades, collecting, organizing, representing and interpreting data are all part of the standards, but these concepts are not described with much richness. In addition, not much attention is given to issues of statistical design.

In the Data Analysis strand, progression is a particular concern from middle to high school. Except for standard 10.D.2, there is little advance in skills and knowledge from grades 7–8 to grades 9–10. At the high school level, one standard in grades 9–10 mentions sample size, but not as part of the larger issue of overall design of surveys and experiments. In comparison to Japan and Arizona, Massachusetts' standards for this strand are not as explicit or rigorous. For example, Japan includes standards for all students for formal counting techniques such as permutations and combinations in grade 10, whereas Massachusetts does not address the topic until grades 11–12. Similarly, Arizona includes expectations about correlation and the use of trend lines to make predictions. Making predictions from data does not appear clearly in the Massachusetts document until grades 11–12 (12.D.3); although it may be implied in grades 9–10 (10.D.2), there does not appear to be any expectation at this level for correlation.

Additionally, there is no specific mention of inferential statistics (e.g., using a density distribution to predict the likelihood that an event has occurred by chance or creating or interpreting a confidence interval). Many of the common mathematical concepts that one finds mentioned in daily newspapers (e.g., indices of inflation or of financial markets, the normal curve, accuracy

limits on polls, the concept of statistical significance, understanding risks) are not in the Pre-K–10 Massachusetts standards required of all students. Such common uses of mathematics are required for informed citizenship and should be included in future editions of the standards.

3. Although the standards generally contain appropriate content, they could be more sharply focused to signal the most essential content at each grade band and to encourage depth.

While the Pre-K–12 Massachusetts standards are comprehensive, they could benefit from a sharpened focus and greater depth. The state rightly targets number sense and operations in the early grades, but many topics within this strand are carried too far into the upper elementary years and even into middle school. Downshifting some of these topics from the middle to the elementary grades would be a move in the right direction. Massachusetts might also wish to include some of the more rigorous aspects of the Japanese standards, with the goal of deepening conceptual understanding. The Japanese elementary school standards for Number Sense tend to set higher and deeper expectations in such areas as representing and ordering fractions and decimals, computations with fractions and decimals, and whole number division.

As it stands now, standards in Number Sense and Operations compete for emphasis with standards in Algebra, Geometry and Data Analysis that should be receiving greater attention as students advance to the middle grades. As a result, the standards in these critical areas are not sufficiently developed. To be sure, Massachusetts has made progress in trying to limit the attention given to arithmetic, as the TIMSS (Third International Mathematics and Science Study) findings strongly suggest is essential, and this is an area that will undoubtedly shift as students and teachers become more knowledgeable about increased expectations in mathematics.

The Geometry strand is another case where the standards cover many ideas at too superficial a level. The depth to which these ideas are developed is not always apparent. For example, in grades 3–4, students are to "identify angles as acute, right or obtuse" and "describe and draw intersecting, parallel and perpendicular lines." The wording of the standard appears to favor knowledge at a vocabulary level rather than deeper understanding. Euclidean geometry is similarly not given adequate play in the middle grades, where the geometry is mostly focused on measurement and vocabulary. Again, the issue appears to be one of depth as opposed to content coverage. Japan is a benchmark to emulate in this regard, as its standards define fewer topics per grade level but go into great depth with those particular topics. Japan has made tough choices about areas of focus for each grade level and defined its choices, albeit without the level of specificity Massachusetts uses.

4. While the introductory statements about the appropriate use of technology are strong, more descriptions within the standards are needed to exemplify how technology can be used as a tool to enhance student learning.

In Guiding Principle III, Massachusetts makes a forthright statement regarding the use of technology in the mathematics classroom. While alluding to the general merits of technology use

in the classroom, the document clearly communicates the philosophy that technology should not be used to replace basic mathematical understanding and skills. Rather, if used properly, it is viewed as a means of enhancing the mathematics curriculum and the learning environment. This is consistent with NCTM's Principles and Standards for School Mathematics, which makes the same point.

With respect to the latter point, Massachusetts might wish to specify what appropriate use of technology looks like in the classroom by identifying opportunities where technology could be put to good use. For example, dynamic geometry software could be used to suggest conjectures or generate perspective drawings, nets and projections. A statistics software package could be used to sample data or create data plots. Computation and graphing technology could be beneficial when learning about compounding. Clarification in the wording of the standards themselves would help, as would the inclusion of sample problems that include appropriately employed technology.

RECOMMENDATIONS FOR MOVING FORWARD

Massachusetts has made substantial progress in developing and implementing two essential components of standards-based reform — strong standards and assessments that measure what the standards expect. As the state moves forward in implementing its system for standards-based education, Achieve recommends that the state consider two improvements in particular to the standards and assessments:

- ✓ **The MCAS high school tests include relatively minor flaws that should be fixed in subsequent rounds of testing. For example, the math test emphasizes some standards at the expense of others, and the English language arts test needs more treatment of nonfiction.** The MCAS mathematics test measures important content that all high school students should be responsible for knowing, yet it can be improved further. Achieve found a number of mathematics test items ostensibly designed to assess number concepts that instead more directly measure algebra standards; this has the effect of weighting the test more heavily toward algebra and omitting some advanced number concepts. Also, while the test is generally well constructed, the balance of items does not provide adequate coverage for all of the important knowledge and skills detailed in the standards. The state should ensure that items assessing numbers and data analysis are appropriately challenging on the next edition of the MCAS mathematics test.

  The language arts test is rich and rigorous, but it focuses too much on literature. All students should be exposed to a deep and engaging literary curriculum, but they also should learn how to read and interpret informational texts, such as historical documents, scientific journal articles and technical manuals. And, the state should consider varying the 10th-grade writing prompt from year to year to assess students' skills in producing different kinds of compositions, rather than just literary analyses. By giving short shrift to students' abilities to work with informational texts, the test does not evaluate the full range of skills students need to participate meaningfully in the emerging "knowledge economy."

- ✓ **The mathematics standards should require more rigor and depth, attention to and emphasis on mathematical reasoning, and a sharper focus on essential content at each grade level.** At the middle and high school levels, students will need to be held to higher expectations to be fully prepared for success in college and high-performance workplaces. While Achieve acknowledges that many Massachusetts schools and students are struggling to meet the state's current standards — and these students will need intensive academic support in the short run — over time, as long-term changes in teacher recruitment, preparation and professional development begin to take effect, the state should raise the rigor of the mathematics standards. In the immediate future, the state may wish to publish companion materials to the mathematics standards that include numerous sample problems and activities

and descriptions of how educators can build students' conceptual understanding of mathematics and reasoning ability.

* * *

Achieve also recently conducted a policy review for Massachusetts, the results of which are planned to be released in late fall of 2001. In the policy review, Achieve will address the findings of this report in addition to other core elements of a comprehensive system of standards-based education, such as capacity building, accountability and public engagement. By taking a hard look at the progress that has been made in the more than eight years Massachusetts has invested in implementing standards-based systems — and by identifying important work still to be done — Achieve hopes to help the state meet its goal of raising the achievement of all its students.

APPENDIX

Achieve relied on a number of individuals in producing this report for Massachusetts and on the expertise of nationally respected experts in academic content, standards, curriculum and assessment design to inform and conduct the standards benchmarking and alignment of assessments to standards.

REPORT CREDITS

Massachusetts Department of Education Project Directors
Jeffrey Nellhaus, Associate Commissioner of Student Assessments
Alan Safran, Senior Associate Commissioner for Student Achievement
Sandra Stotsky, Senior Associate Commissioner, Office of Standards, Academic Review and Research

Achieve Project Directors and Report Authors
Jean B. Slattery, Associate Director, Achieve's Benchmarking Initiative
Jennifer L. Vranek, Director, Benchmarking and State Services, Achieve, Inc.

Achieve Senior Project Staff
Karin Elsea, Project Administrator
JoAnne T. Eresh, Senior Associate for English language arts, Achieve's Benchmarking Initiative
Kaye R. Forgione, Senior Associate for Mathematics, Achieve's Benchmarking Initiative

Achieve Advisors
Matthew D. Gandal, Vice President and Director of the Washington Office, Achieve, Inc.
Lauren B. Resnick, Director, Learning Research and Development Center, University of Pittsburgh
Robert B. Schwartz, President, Achieve, Inc.

Report Production
Editorial and design assistance was provided by Kathy Delaney, Beth Rennie and Andrea Sussman of KSA-Plus Communications.

EXPERT AND STAFF BIOGRAPHIES

ENGLISH LANGUAGE ARTS

JoAnne Thibault Eresh is a senior associate at Achieve, where she leads the English language arts aspects of the standards-to-standards benchmarking and assessment-to-standards alignment reviews. She taught writing at the university level and English at private and public high schools in St. Louis, Mo., and in Fitchburg, Mass. She began her work in curriculum design and performance assessment in 1979 under Superintendent Richard C. Wallace, Jr., and from 1981 to 1994, she was director of the division of writing and speaking for the Pittsburgh Public Schools. During that time

she directed The Pittsburgh Discussion Model Project, funded by the Rockefeller Foundation and part of the CHART network, and later directed the imaginative writing part of the ARTS Propel Project, a joint project with Harvard's Project Zero and the Educational Testing Service. She was the Pittsburgh district coordinator for the New Standards Project and wrote the teachers' guides for the New Standards ELA Portfolios. In 1995, she was one of the original resident fellows at the Institute for Learning at the University of Pittsburgh's Learning Research and Development Center and coordinated the New Standards Linking Projects. From 1997 to March 2001, she was the coordinator of staff development in Community District Two in New York City, where she was responsible for the hiring, training and coordination of that district's staff development group.

Assessment Reviewers

Carlton Jordan received his mater's degree from the State University of New York at Albany. Mr. Jordan is a senior associate with The Education Trust, where his primary responsibilities include helping K–12 teachers implement standards-based education and helping middle and high school English language arts departments de-track English language arts. Prior to joining the Trust in August 1998, Mr. Jordan was an adjunct lecturer in the department of African Studies at the State University of New York at Albany; previously, he taught middle and high school English language arts in Montclair, N.J. The efforts of Mr. Jordan and his colleagues in de-tracking 9th-grade English language arts led to an article by Michelle Fine, Lois Weiss and Linda Powell in a book called *Critical Ethnicity* and a video called *Off Track: Classroom Privilege for All* distributed by Teachers College Press. The video won the National Educational Media Network's Silver Apple award.

Laura McGiffert is the assistant director of the Mathematics Achievement Partnership (MAP), a multistate collaboration to dramatically improve mathematics performance in the middle grades. As the principal manager of this project, Ms. McGiffert assumed primary responsibility for the development of *Foundations for Success: Mathematics for the Middle Grades*, which represents the core knowledge and skills that students should learn to prepared for high school and beyond. To this end, she coordinated Achieve's Mathematics Advisory Panel, an expert panel of mathematicians, mathematics educators, curriculum specialists, and state and local math supervisors representing a broad spectrum of perspectives about mathematics education. Before joining Achieve in 1998, Ms. McGiffert was a high school English language arts teacher for five years in Colorado, where she was involved in a districtwide effort to refine and align local standards and assessments. In 1995, she was awarded "Best First Year Teacher" in Eagle County School District at the high school level. She also taught writing and literature at Colorado Mountain College. Currently, she brings these experiences to bear as a member of the English language arts team for Achieve's Benchmarking Initiative, which assists states in benchmarking their academic standards, assessments and accountability policies against the best in the nation and the world. A native Washingtonian, Ms. McGiffert hold a master's degree in education policy from the Georgetown University Public Policy Institute, a master's degree in secondary education from the University of Colorado at Boulder and a bachelor's degree in English and American literature from Harvard University.

LeRoy Miller received his bachelor's degree in English from Fairleigh Dickinson University and his master's in Secondary English from Miami University. Mr. Miller is in his 28th year of teaching 11th-grade English at Sidney High School, where he also serves as English Department Chair for a nine-member department. Mr. Miller is also a teacher of advanced placement and honors American literature and an advisor for the Academic Competition Team. In addition, he served as a commissioner on the Governor's Commission on Student Success in 2000 and is a member of Ohio's Academic Content Standards Writing Team for English language arts.

Diane Ravitch received her doctorate from Columbia University. Dr. Ravitch is a research professor at New York University and holds the Brown Chair in Education Policy at the Brookings Institution in Washington, D.C., where she is a senior fellow and edits the *Brookings Papers on Education Policy*. She is also a member of the National Assessment Governing Board, to which Secretary of Education Riley appointed her in 1997. Previously, Dr. Ravitch served as assistant secretary for educational research and improvement in the George H.W. Bush Administration. The author of numerous books, scholarly and popular essays, she most recently published *Left Back: A Century of Failed School Reform* (2000).

Alice Turner Venson received her bachelor's degree in English from Hampton University (formerly Hampton Institute in Virginia), a master's degree in education from the University of Pittsburgh, and has done doctoral study in reading and language arts K–12 at the University of Pittsburgh. Ms. Turner Venson is a senior associate with the National Center on Education and the Economy. Her area of concentration is English language arts at the secondary level. Previously, she coordinated the New American Schools school reform initiative for the Pittsburgh Public Schools. Ms. Turner Venson has also been a high school English teacher, guidance counselor, vice principal, instructional supervisory specialist for English at the secondary level, a state and national trainer for writing assessment, and a national trainer for standards development.

Mathematics

Kaye R. Forgione began consulting work with Achieve in 2000 and joined Achieve as senior associate for mathematics in March 2001. Dr. Forgione's primary responsibilities are managing and providing intellectual leadership to Achieve's standards and benchmarking work involving mathematics. Before joining Achieve, she served as assistant director of the Systemic Research Collaborative for Mathematics, Science and Technology Education (SYRCE) project at the University of Texas at Austin. Her responsibilities also included administrative and design responsibilities for UTeach, a collaborative project of the College of Education and the College of Natural Sciences to train and support the next generation of math and science teachers in Texas. Prior to her work at the University of Texas, Dr. Forgione was director of academic standards programs at the Council for Basic Education, a nonprofit education organization based in Washington, D.C. Prior to joining the Council for Basic Education in 1997, Dr. Forgione worked in the K–12 arena in a variety of contexts, including district-level curriculum supervisor for mathematics, assessment and gifted/talented programs. She was also team leader for assessments and task development and supervisor in the areas of assessment, school profiles, and educational

indicators at the Delaware Department of Education, senior research associate for development at the Delaware Education Research and Development Center at the University of Delaware, and a high school mathematics teacher. Her personal portfolio of work includes mathematics-related and policy development work in a number of states (including Nevada, Maryland and Hawaii), school districts (including Cleveland; East Allen County, Ind.; and Los Angeles) and in partnership with a number of organizations (including Achieve, George Washington University, and the Institute for Educational Leadership). Dr. Forgione earned her doctorate from the University of Delaware.

Standards Reviewers

Susan K. Eddins has both a bachelor's and master's degree in mathematics. She has taught students in kindergarten through college for over 30 years — 24 years at the high school level. She is the recipient of several honors for her teaching, including the Presidential Award for Excellence in Mathematics Teaching, and she is a National Board Certified Teacher in Adolescent and Young Adult Mathematics. Ms. Eddins is a faculty member, an instructional facilitator and the curriculum and assessment leader in mathematics at the Illinois Mathematics and Science Academy, where she has taught since the school's inception in 1986. She has served in leadership capacities in several professional organizations and is currently a member of the Board of Directors of the National Council of Teachers of Mathematics (NCTM). Ms. Eddins was a member of the 9–12 writing group for NCTM's *Principles and Standards for School Mathematics.* Ms. Eddins is co-author of a chapter in NCTM's *Windows of Opportunity* and is a co-author of *UCSMP Algebra.* Ms. Eddins is a past panel member and editor of NCTM's *Student Math Notes* and has authored several articles in refereed journals. Over the past four years, in addition to numerous workshops and presentations, her most extensive work has been in the area of standards development, standards review and alignment of standards to assessments. For Achieve, she has reviewed academic standards or assessments from Alaska, Illinois, Indiana, Minnesota, New Jersey, Oregon, Pennsylvania, Texas and Washington.

Roger E. Howe received his doctorate in mathematics from the University of California at Berkeley. Dr. Howe currently is a professor of mathematics at Yale University. He is a member of the National Academy of Sciences and the American Academy of Arts and Sciences. Dr. Howe is a member of the Board of Directors, Connecticut Academy for Education in Mathematics, Science and Technology. He was the chair of American Mathematical Society Consultative Committee to the NCTM Standards 2000, served on the Mathematical Association of America Committee on Education of Mathematics Teachers and participated in the National Research Council Study Committee on Mathematics Learning.

Donald R. King received his doctorate in mathematics from the Massachusetts Institute of Technology. Dr. King is an associate professor of Mathematics at Northeastern University. Previously, he was a visiting assistant professor at Salem State College, a visiting assistant professor at the University of California at San Diego and a high school mathematics teacher in Boston, Mass. Dr. King is a member of the Mathematical Association of America, American Mathematical Society and the National Association of Mathematicians. Dr. King is active in professional and community service: He was a parent member of the Mathematics Focus Group

for Boston Public Schools in 1997; director from 1993 to 1994 of NUMATH, Northeastern University's program to foster minority mathematical achievement and talent in high school; an advisor to algebra-in-middle-schools projects from 1990 to 1992; a review panelist for three years for Ford Foundation post-doctoral fellowships for minorities; and an advisor to Massachusetts' pre-engineering program for minorities from 1988 to 1991. Dr. King recently gave a speech at the American Mathematical Society's Special Session on Teaching Mathematics in the New Millennium: "Changing school outcomes: Raising standards and promoting equity," and he has advised Achieve on the quality of standards and assessments in a number of states including Minnesota, Oklahoma, New Jersey and Texas.

Lynn A. Steen received his doctorate in mathematics from the Massachusetts Institute of Technology. Dr. Steen has been a member of the St. Olaf College faculty since 1965 and currently is the director of institutional research and planning. Dr. Steen is leader of the quantitative literacy initiative of the National Council on Education and the Disciplines (NCED) at the Woodrow Wilson Foundation and a member of the Mathematics Achievement Partnership (MAP), a project of Achieve, Inc., to create a common set of expectations and assessments for mathematics in the middle grades. Dr. Steen is former president of the Mathematical Association of America and former chair of the Council of Scientific Society Presidents. He has reviewed mathematics standards for Achieve from Texas, Oklahoma and Massachusetts.

Assessment Reviewers

Lucio Calzada received his master's in school administration from Texas A&I University and is an advanced doctoral candidate at the Texas A&M University. Mr. Calzada is currently the principal of Fulmore Middle School in Austin, Texas. Prior to joining Fulmore, Mr. Calzada was the principal of South Park Middle School in Corpus Christi, Texas, where he played a role in the development and the implementation of the academic standards supported by the Edna McConnell Clark Foundation. He has also been a mathematics consultant, an assistant principal of Taft Junior High School, and teacher of mathematics, trigonometry, analytic geometry, geometry, algebra and prealgebra at H.M. King High School in Kingsville, Texas.

Ruben Carriedo received his doctorate from the Harvard University Graduate School of Education. Dr. Carriedo currently serves as a senior research associate at the University of Michigan School of Education. He is part of the Study of Instructional Improvement, a comprehensive longitudinal study of instructional interventions at 125 elementary schools across the country. Dr. Carriedo has been a teacher, coach, counselor, and secondary school and central office administrator in New York City and San Diego public schools. He served as the assistant superintendent for the Planning Assessment, Accountability and Development Division in the San Diego City Schools from 1987 to 1999. He has served on the California Assessment Program Policy and Technical Advisory committees, New Standards Project Technical Committee, the National Advisory Panel for the Center on the Reorganization and Restructuring of Schools, the National Advisory Panel on Chicago School Restructuring, PACE/Rockefeller Project Advisory Committee, *Editorial Board of the Educational Evaluation and Policy Analysis Journal,* Advisory Committee of the Learning, Research and Development Center at the University of Pittsburgh. He

currently serves on the TIMSS-R Technical Review Panel, Advisory Committee on Education Statistics at the National Center for Education Statistics, Trustee Committee on Research and Development of the College Board, and Title I Assessment Committee at the National Research Council and National Academy of Science. He has worked as a consultant to the Spencer and MacArthur foundations and Booz-Allen on Chicago school reform. In addition, he has consulted with the Ford Foundation to plan a national symposium on assessment and equity. He also served as a member of the national evaluation panel for the Children Achieving reform initiative in Philadelphia.

Donald R. King received his doctorate in mathematics from the Massachusetts Institute of Technology. Dr. King is an associate professor of Mathematics at Northeastern University. Previously, he was a visiting assistant professor at Salem State College, a visiting assistant professor at the University of California at San Diego and a high school mathematics teacher in Boston, Mass. Dr. King is a member of the Mathematical Association of America, American Mathematical Society and the National Association of Mathematicians. Dr. King is active in professional and community service: He was a parent member of the Mathematics Focus Group for Boston Public Schools in 1997; director from 1993 to 1994 of NUMATH, Northeastern University's program to foster minority mathematical achievement and talent in high school; an advisor to algebra-in-middle-schools projects from 1990 to 1992; a review panelist for three years for Ford Foundation post-doctoral fellowships for minorities; and an advisor to Massachusetts' pre-engineering program for minorities from 1988 to 1991. Dr. King recently gave a speech at the American Mathematical Society's Special Session on Teaching Mathematics in the New Millennium: "Changing school outcomes: Raising standards and promoting equity," and he has advised Achieve on the quality of standards and assessments in a number of states including Minnesota, Oklahoma, New Jersey and Texas.

Norman Webb received his doctorate from Stanford University. Dr. Webb is a senior research scientist at the Wisconsin Center for Education Research, located in the University of Wisconsin at Madison. Before joining the University of Wisconsin, Dr. Webb worked for Wisconsin's Department of Public Instruction, as section chief for the Bureau of Achievement Testing. Dr. Webb was recently awarded a grant by the Joyce Foundation entitled Center for the Study of Systemic Reform in Milwaukee Public Schools. In addition, Dr. Webb authored *Alignment of Science and Mathematics Standards and Assessments in Four States* for the National Institute for Science Education and the Council of Chief State School Officers in 1999. He has worked with Achieve to review tests for Illinois, Minnesota, New Jersey and Pennsylvania and to refine Achieve's assessment-to-standards alignment protocol.

Achieve Project Staff

Matthew D. Gandal is vice president of Achieve, Inc. Mr. Gandal manages the Washington, D.C., office and is responsible for overseeing Achieve's major initiatives. These include the 1999 National Education Summit and a series of follow-up activities Achieve has launched to help states address the Summit challenges; the Benchmarking Initiative that helps states compare their

standards, assessments and accountability policies with other states and nations; and the Mathematics Achievement Partnership, designed to help states improve curriculum and instruction in middle school mathematics and measure student achievement using a common, internationally benchmarked 8th-grade test. Before joining Achieve, Mr. Gandal was assistant director for educational issues at the American Federation of Teachers. Mr. Gandal helped the AFT launch a variety of programs and publications designed to support standards-based reform efforts in states and school districts. He was the author and chief architect of *Making Standards Matter,* an annual AFT report evaluating the quality of the academic standards, assessments and accountability policies in the 50 states. He also authored a series of reports that compared student standards and achievement in the United States with that of other industrialized nations. Mr. Gandal is a proud graduate of the public school system in the state of Maryland and earned a bachelor's degree in philosophy from Trinity College in Hartford, Conn.

Lauren B. Resnick is an internationally known scholar in the cognitive science of learning and instruction. Her recent research has focused on socializing intelligence, the nature and development of thinking abilities, and the relation between school learning and everyday competence, with special attention to mathematics and literacy. Her current work lies at the intersection of cognitive science and policy for education. Dr. Resnick founded and directs the Institute for Learning, which focuses on professional development based on cognitive learning principles and effort-oriented education. She is cofounder and codirector of the New Standards Project, which has developed standards and assessments that have widely influenced state and school district practice. Dr. Resnick was a member of the Commission on the Skills of the American Workforce and served as chair of the assessment committee of the SCANS Commission and of the Resource Group on Student Achievement of the National Education Goals Panel. She has served on the Commission on Behavioral and Social Sciences and Education and on the Mathematical Sciences Education Board at the National Research Council. Her National Academy of Sciences monograph, *Education and Learning to Think,* has been influential in school reform efforts, and her widely circulated Presidential Address to the American Educational Research Association, "Learning In School and Out," has shaped thinking about youth apprenticeship and school-to-work transition. Dr. Resnick is professor of psychology at the University of Pittsburgh, where she directs the prestigious Learning Research and Development Center. Educated at Radcliff and Harvard, she has been a member of the Smithsonian Council and was the recipient of the 1998 E.L.Thorndike Award from the American Psychological Association and the 1999 Oeuvre Award from the European Association for Research on Learning and Instruction.

Robert Schwartz is the president of Achieve, Inc. Over the previous three-and-a-half decades, Mr. Schwartz has had a rich and varied career in education and government. He has been a high school English teacher and principal; an education advisor to the mayor of Boston and governor of Massachusetts; an assistant director of the National Institute of Education; a special assistant to the president of the University of Massachusetts; the executive director of the Boston Compact, a public-private partnership designed to improve access to higher education and employment for urban high school graduates; and a lecturer on education at the Harvard University Graduate School of Education. From 1990 to 1996, Mr. Schwartz directed the education grant-making

program of The Pew Charitable Trusts, one of the nation's largest private philanthropies. Among the major reform projects initiated during his tenure at the Trusts were New Standards, a voluntary national system of student performance standards and assessments developed jointly by the University of Pittsburgh, the National Center on Education and the Economy, and 17 partner states; and the Pew Network for Standards-Based Reform, a collaborative venture among seven medium-size school districts committed to systemic reform based on high academic standards. Mr. Schwartz has written and spoken widely on such topics as urban school reform, public-private partnerships and the role of higher education in K–12 reform. He holds degrees from Harvard and Brandeis Universities and continues to serve as a part-time faculty member at Harvard, where he teaches a course each spring on educational policy and administration.

Jean B. Slattery has been a consultant for Achieve since 1999; she currently serves as associate director, Benchmarking Initiative. She was supervising director of curriculum development and support in Rochester, N.Y., from 1989 to 1997, with responsibility for overseeing the work of all subject-area directors in the K–12 instructional program. Her earlier responsibilities as a district-level administrator included serving as director of the Middle School (1987–89) and Junior High (1985–87) Programs. During this period, she initiated Teachers As Partners, a peer-coaching staff development program funded by the Ford and Matsushita (Panasonic) Foundations. Prior to her work in central office, Dr. Slattery served as vice principal of the Nathaniel Rochester Community School. She taught chemistry for 15 years in Rochester, New Haven and Branford, Conn. She was district coordinator for the New Standards Project of the National Center on Education and the Economy (NCEE) from 1991 to 1997 and served on the faculty of the July Institutes on Assessment of the Harvard University School of Education from 1990 to 1999. Dr. Slattery also is a peer consultant on standards and assessment for the U.S. Department of Education. She has served as a consultant to the Washington, D.C. School District, the San Diego Unified School District, a Washington state consortium of rural schools, and the Alabama and Illinois Departments of Education. She has worked for the Council for Basic Education on projects involving the Flint Community School District, the Nevada Education Department and the Cleveland Municipal School District. She received a bachelor's degree in chemistry from Albertus Magnus College, a master's in science education from Yale University and a doctorate in science curriculum from the University of Rochester.

Jennifer L. Vranek is the director of Benchmarking and State Services for Achieve, Inc., where she has directed successful projects with more than 15 states, working closely with state education superintendents, governors and business executives to benchmark state education reforms and share best practices. Ms. Vranek was a key staff member in the planning, preparation and follow-up activities for the 1999 and 2001 National Education Summits, hosted by Achieve, and in the research, development and launch of the Mathematics Achievement Partnership. In addition, Ms. Vranek recently planned and launched a two-year, $2.4-million project spearheaded by Achieve in collaboration with The Education Trust, Thomas B. Fordham Foundation and the National Alliance of Business to help states close the gap between current high school graduation standards and college/employer admissions. Before joining Achieve in 1997, Ms. Vranek was a research assistant at the American Federation of Teachers, where she was a principal researcher for *Making*

Standards Matter, an annual AFT report evaluating the quality of the academic standards, assessments and accountability policies in the 50 states, and *Setting Higher Sights,* a comparative analysis of the quality of mathematics assessments in the United States and abroad. Previously, Ms. Vranek lived in Brasilia, Brazil, where she was a project consultant to the World Bank's G-7 Pilot Program to Conserve the Brazilian Rainforest and also served as the administrative secretary to the Board of Directors of the American School of Brasilia. A graduate of the public schools in San Antonio, Texas, Ms. Vranek holds a master's degree in public policy from the Georgetown University Public Policy Institute and a bachelor's degree in history with minors in music and public policy from the College of William and Mary in Virginia.

Board of Directors, Achieve, Inc.

Co-Chairs

John Engler, Governor
State of Michigan

Louis V. Gerstner, Jr., Chairman and CEO
IBM Corporation

Co-Vice Chairs

Philip M. Condit, Chairman and CEO
The Boeing Company

Gray Davis, Governor
State of California

Board Members

Keith E. Bailey, Chairman, President and CEO
Williams

Roy E. Barnes, Governor
State of Georgia

Craig R. Barrett, President and CEO
Intel Corporation

Frank Keating, Governor
State of Oklahoma

Gary Locke, Governor
State of Washington

Edward B. Rust Jr., Chairman and CEO
State Farm Insurance

Arthur F. Ryan, Chairman and CEO
Prudential

Bob Taft, Governor
State of Ohio

President

Robert B. Schwartz

Achieve, Inc.

Achieve, Inc.
www.achieve.org

400 North Capitol Street, NW
Suite 351
Washington, DC 20001
Phone: (202) 624-1460
Fax: (202) 624-1468

8 Story Street
1st Floor
Cambridge, MA 02138
Phone: (617) 496-6300
Fax: (617) 496-6361